Numerical Methods for Computing Angles Between Linear Subspaces

By Åke Björck and Gene H. Golub*

Abstract. Assume that two subspaces F and G of a unitary space are defined as the ranges (or null spaces) of given rectangular matrices A and B. Accurate numerical methods are developed for computing the principal angles $\theta_k(F, G)$ and orthogonal sets of principal vectors $u_k \in F$ and $v_k \in G, k = 1, 2, \cdots, q = \dim(G) \leqq \dim(F)$. An important application in statistics is computing the canonical correlations $\sigma_k = \cos \theta_k$ between two sets of variates. A perturbation analysis shows that the condition number for θ_k essentially is $\max(\kappa(A), \kappa(B))$, where κ denotes the condition number of a matrix. The algorithms are based on a preliminary QR-factorization of A and B (or A^H and B^H), for which either the method of Householder transformations (HT) or the modified Gram-Schmidt method (MGS) is used. Then $\cos \theta_k$ and $\sin \theta_k$ are computed as the singular values of certain related matrices. Experimental results are given, which indicates that MGS gives θ_k with equal precision and fewer arithmetic operations than HT. However, HT gives principal vectors, which are orthogonal to working accuracy, which is not generally true for MGS. Finally, the case when A and/or B are rank deficient is discussed.

1. Introduction. Let F and G be given subspaces of a unitary space E^m, and assume that

$$(1) \qquad p = \dim(F) \geqq \dim(G) = q \geqq 1.$$

The smallest angle $\theta_1(F, G) = \theta_1 \in [0, \pi/2]$ between F and G is defined by

$$\cos \theta_1 = \max_{u \in F} \max_{v \in G} u^H v, \qquad ||u||_2 = 1, ||v||_2 = 1.$$

Assume that the maximum is attained for $u = u_1$ and $v = v_1$. Then, $\theta_2(F, G)$ is defined as the smallest angle between the orthogonal complement of F with respect to u_1 and that of G with respect to v_1. Continuing in this way until one of the subspaces is empty, we are led to the following definition.

Definition. The principal angles $\theta_k \in [0, \pi/2]$ between F and G are recursively defined for $k = 1, 2, \cdots, q$ by

$$(2) \qquad \cos \theta_k = \max_{u \in F} \max_{v \in G} u^H v = u_k^H v_k, \qquad ||u||_2 = 1, ||v||_2 = 1,$$

subject to the constraints

$$u_j^H u = 0, \qquad v_j^H v = 0, \qquad j = 1, 2, \cdots, k - 1.$$

The vectors $(u_1, \cdots, u_q)$ and $(v_1, \cdots, v_q)$ are called principal vectors of the pair of spaces.

We note that the principal vectors need not be uniquely defined, but the principal angles always are. The vectors $V = (v_1, \cdots, v_q)$ form a unitary basis for G and the vectors $U = (u_1, \cdots, u_q)$ can be complemented with $(p - q)$ unitary vectors so that $(u_1, \cdots, u_p)$ form a unitary basis for F. It can also be shown that

$$u_j^H v_k = 0, \qquad j \neq k, j = 1, \cdots, p, k = 1, \cdots, q.$$

For an introduction to these concepts, we refer to [1]. An up to date list of references can be found in [9].

Principal angles and vectors have many important applications in statistics and numerical analysis. In [7], the statistical models of canonical correlations, factor analysis and stochastic equations are described in these terms. The eigenvalue problem $Ax = \lambda Bx$ can have continuous eigenvalues if the nullspaces associated with A and B intersect [13]. By taking the vectors u_k corresponding to $\cos \theta_k = 1$, we get a unitary basis for the intersection, which can be used to simultaneously deflate A and B. Other applications are found in the theory of approximate least squares [8] and in the computation of invariant subspaces of a matrix [21].

The purpose of this paper is to develop new and more accurate methods for computing principal angles and vectors, when the subspaces are defined as the ranges (or nullspaces) of two given matrices A and B. In Section 2, we describe the standard method of computing canonical correlations and show why this method may give rise to a serious loss of accuracy. Assuming that unitary bases for F and G are known, we derive, in Section 3, formulas for computing principal angles and vectors from the singular values and vectors for certain matrices. To find out how accurately the angles are defined in the presence of uncertainties in A and B, first order perturbation results are given in Section 4. In Section 5, different numerical methods for computing the unitary bases, and the use of the formulas from Section 3, are discussed with respect to efficiency and accuracy. The special problems which arise when A and/or B are exactly or nearly rank deficient are discussed in Section 6. Finally, some numerical results are given in Section 7.

2. Canonical Correlations.

For a matrix A, we denote the range of A by $R(A)$ and the nullspace of A by $N(A)$:

$$(3) \qquad R(A) = \{u \mid Ax = u\}, \qquad N(A) = \{x \mid Ax = 0\}.$$

In the problem of canonical correlations, we have $F = R(A)$, $G = R(B)$ where A and B are given rectangular matrices. Then, the canonical correlations are equal to $\cos \theta_k$, and it can be shown that

$$(4) \qquad \cos \theta_k = \sigma_k, \qquad u_k = Ay_k, \qquad v_k = Bz_k, \qquad k = 1, 2, \cdots, q,$$

where $\sigma_k \geqq 0$ are eigenvalues and y_k, z_k properly normalized eigenvectors to the generalized eigenvalue problem

$$(5) \qquad \begin{bmatrix} 0 & A^H B \\ B^H A & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \sigma \begin{bmatrix} A^H A & 0 \\ 0 & B^H B \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix}.$$

Assume for convenience that A and B have full column rank. The standard method [6] of computing canonical correlations is to compute $A^H A$, $B^H B$, $A^H B$ and perform the Choleski decompositions

$$A^H A = R_A^H R_A, \qquad B^H B = R_B^H R_B,$$

where R_A and R_B are upper triangular.

The eigenvalue problem (5) is then equivalent to the eigenvalue problems for a pair of Hermitian matrices

$$M M^H \hat{y}_i = \sigma_i^2 \hat{y}_i, \qquad M^H M \hat{z}_i = \sigma_i^2 \hat{z}_i$$

where

$$M = (R_A^H)^{-1} (A^H B) R_B^{-1}, \qquad \hat{y}_i = R_A y_i, \qquad \hat{z}_i = R_B z_i.$$

These can be solved by standard numerical methods.

When $q = 1$ and $B = b$, the principal angles and vectors are closely related to the least squares problem of minimizing $\|b - Ax\|_2$. In fact, with the notations above (but dropping subscripts), we have

$$y = x/\|Ax\|_2, \qquad z = 1/\|b\|_2, \qquad \sigma = \|Ax\|_2/\|b\|_2,$$

and (5) is reduced to

$$A^H bz = \sigma A^H Ay, \qquad b^H Ay = \sigma b^H bz.$$

But the first equation here is the normal equations for $x = \sigma y/z$. Thus, the classical algorithm reduces for $q = 1$ to solution of the normal equations by Choleski's method.

Lately it has been stressed by several authors that forming the normal equations in single precision involves a loss of information which cannot be retrieved. For linear least squares problems, other methods without this disadvantage have been developed ([2], [16] and [17]). Our aim in this paper is to generalize these methods to the case when $q > 1$.

3. Solution Using Singular Values. In most applications, each subspace is defined as the range, or the complement of the range, of a given matrix. In this case, a unitary basis for the subspace may be computed in a numerically stable way by well-known methods for the QR-decomposition of a matrix. These methods will produce for an $m \times n$ matrix A, with $m \geq n$, a decomposition

$$A = (Q' \mid Q'') \binom{S}{0} \begin{matrix} \}p \times n \\ \}(m - p) \times n \end{matrix}$$

where rank$(S) = p$ and $Q = (Q' \mid Q'')$ is unitary. Then Q' gives a unitary basis for the range of A, $R(A)$, and Q'' a unitary basis for the complement $[R(A)]^-$. Notice that the case when a subspace is defined as the nullspace $N(A^H)$ of a matrix A^H is included, since $N(A^H) = [R(A)]^-$. The computation of unitary bases will be discussed in more detail in Sections 5 and 6, and we assume here that such bases have been obtained.

Recently, an efficient and numerically stable algorithm for computing the singular value decomposition [11] (SVD) of a matrix has been developed [17]. This algorithm

will be our basic tool for computing principal angles and vectors. The relation between singular values and our problem is clear from the following theorem.

THEOREM 1. *Assume that the columns of Q_A and Q_B form unitary bases for two subspaces of a unitary space E^m. Put*

$$(6) \qquad M = Q_A^H Q_B,$$

and let the SVD of this $p \times q$ matrix be

$$(7) \qquad M = YCZ^H, \qquad C = \mathrm{diag}(\sigma_1, \cdots, \sigma_q),$$

where $Y^H Y = Z^H Z = ZZ^H = I_q$. If we assume that $\sigma_1 \geqq \sigma_2 \geqq \cdots \geqq \sigma_q$, then the principal angles and principal vectors associated with this pair of subspaces are given by

$$(8) \qquad \cos \theta_k = \sigma_k(M), \qquad U = Q_A Y, \qquad V = Q_B Z.$$

Proof. It is known [18] that the singular values and singular vectors of a matrix M can be characterized by

$$(9) \qquad \sigma_k = \max_{\|y\|_2 = \|z\|_2 = 1} (y^H M z) = y_k^H M z_k,$$

subject to

$$y^H y_j = z^H z_j = 0, \qquad j = 1, \cdots, k - 1.$$

If we put

$$u = Q_A y \in R(Q_A), \qquad v = Q_B z \in R(Q_B),$$

then it follows that $\|u\|_2 = \|y\|_2$, $\|v\|_2 = \|z\|_2$ and

$$y^H y_j = u^H u_j, \qquad z^H z_j = v^H v_j.$$

Since $y^H M z = y^H Q_A^H Q_B z = u^H v$, (9) is equivalent to

$$\sigma_k = \max_{\|u\|_2 = \|v\|_2 = 1} (u^H v) = u_k^H v_k$$

subject to

$$u^H u_j = v^H v_j = 0, \qquad j = 1, \cdots, k - 1.$$

Now (8) follows directly from the definition of principal angles and vectors (2), which concludes the proof.

For small angles, θ_k is not well determined from $\cos \theta_k$ and we now develop formulas for computing $\sin \theta_k$. Let Q_A and Q_B be defined as in Theorem 1. For convenience, we change the notations slightly and write (7) and (8) as

$$(10) \qquad M = Y_A C Y_B^H, \qquad U_A = Q_A Y_A, \qquad U_B = Q_B Y_B.$$

We split Q_B according to

$$(11) \qquad Q_B = P_A Q_B + (I - P_A) Q_B = P_A Q_B + P_A^\perp Q_B,$$

where $P_A = Q_A Q_A^H$ is the orthogonal projection onto $R(Q_A)$. Here

$$P_A Q_B = Q_A Q_A^H Q_B = Q_A M = Q_A Y_A C Y_B^H,$$

and hence the SVD of the matrix $P_A Q_B$ is given by

(12) $$P_A Q_B = U_A C Y_B^H, \qquad C = \text{diag}(\cos \theta_k).$$

Since $P_A(I - P_A) = 0$, we get, from squaring (11),

$$Q_B^H(I - P_A)^2 Q_B = I - Q_B^H P_A^2 Q_B = Y_B(I - C^2) Y_B^H.$$

It follows that the SVD of $(I - P_A)Q_B$ can be written

(13) $$(I - P_A)Q_B = W_A S Y_B^H, \qquad S = \text{diag}(\sin \theta_k).$$

Comparing (13) with (12), it is evident that W_A gives the principal vectors in the complement $[R(Q_A)]^-$ associated with the pair of subspaces $([R(Q_A)]^-, R(Q_B))$.

We will, for the rest of this section, assume that, in addition to (1), we have $p + q \leqq m$. (This is no real restriction, since, otherwise, we have $(m - p) + (m - q) \leqq m$, and we can work with the complements of $R(Q_A)$ and $R(Q_B)$ instead.) Then, $\dim([R(Q_A)]^-) = m - p \geqq q$, and we can choose the $m \times q$ matrix W_A in (13) so that $W_A^H U_A = 0$.

By analogy, we have formulas similar to (12) and (13) related to the splitting $Q_A = P_B Q_A + (I - P_B)Q_A$,

(14) $$P_B Q_A = U_B C Y_A^H, \qquad (I - P_B)Q_A = W_B S Y_A^H,$$

where again, since $m - q \geqq p \geqq q$, we can choose the $m \times q$ matrix W_B so that $W_B^H U_B^H = 0$. From (14), we get

$$U_A = Q_A Y_A = (U_B C + W_B S) Y_A^H Y_A = (U_B \ W_B) \binom{C}{S}.$$

If we put

$$P_{B,A} = U_A U_B^H = (U_B \ W_B) \binom{C}{S} U_B^H,$$

then, since $R(Q_B) = R(U_B)$, we have for any $x \in R(Q_B)$ that $P_{B,A} x = U_A U_B^H(U_B y) = U_A y$, and thus

$$P_{B,A} x \in R(Q_A), \qquad ||x||_2 = ||P_{B,A} x||_2.$$

We can now always find an $m \times (m - 2q)$ matrix Z_B such that $(U_B W_B Z_B)$ is a unitary basis in E^m. Then

(15) $$P_{B,A} = (U_B W_B \mid Z_B) \begin{bmatrix} C & -S & 0 \\ S & C & \\ \hline & & \\ 0 & & I \end{bmatrix} \begin{bmatrix} U_B^H \\ W_B^H \\ Z_B^H \end{bmatrix}$$

is the matrix of a unitary transformation, mapping $R(Q_B)$ into $R(Q_A)$. Its restriction to $R(Q_B)$ is $P_{B,A}$, and it leaves all vectors in $R(Z_B)$ unchanged. This transformation is called a *direct rotation* [9] from $R(Q_B)$ into $R(Q_A)$. It is distinguished from other unitary transformations P taking $R(Q_B)$ into $R(Q_A)$ by the property that it minimizes each unitarily invariant norm of $(I - P)^H(I - P)$. If $R(Q_B) \cap [R(Q_A)]^-$ is empty, then all $\theta_k < \pi/2$ and the direct rotation is uniquely determined.

Similarly, we can construct a direct rotation taking $R(U_A)$ into $R(Q_B)$. It is obvious that the relations between the two subspaces are very completely characterized by the quantities C, S, U_A, W_A, U_B and W_B.

4. Perturbation of Principal Angles. We consider here how the principal angles between $R(A)$ and $R(B)$ change when the elements in A and B are subject to perturbations. We assume in this analysis that the matrices A and B are $m \times p$ and $m \times q$, respectively, and have linearly independent columns. Consider first a perturbation of A only,

$$A_\epsilon = A + \epsilon E = (A + \epsilon E_1) + \epsilon E_2,$$

where we have split the perturbation into components in and orthogonal to $R(A)$,

$$E_1 = P_A E, \qquad E_2 = (I - P_A) E.$$

Let the polar decomposition of $A + \epsilon E_1$ be

$$A + \epsilon E_1 = Q_A H_A, \qquad Q_A^H Q_A = I, \qquad H_A \text{ positive definite}.$$

Then, since $R(A) = R(A + \epsilon E_1)$, Q_A gives a unitary basis for $R(A)$.

To get a unitary basis for $R(A_\epsilon)$, we first note that for small absolute values of ϵ, the matrix

$$(A + \epsilon E) H_A^{-1} = Q_A + \epsilon F_2, \qquad F_2 = (I - P_A) F, \qquad F = E H_A^{-1},$$

is nearly orthogonal. Indeed, since $Q_A^H F_2 = Q_A^H P_A F_2 = 0$, we have

$$S = I - (Q_A + \epsilon F_2)^H (Q_A + \epsilon F_2) = -\epsilon^2 F_2^H F_2,$$

and

(16) $$\sigma_1(S) = \epsilon^2 \cdot \sigma_1^2(F_2) \leqq \epsilon^2 \sigma_1^2(F).$$

Then, using a series expansion from [4] for the unitary factor Q_{A_ϵ} in the polar decomposition of $A_\epsilon H_A^{-1}$, it follows that

$$(Q_A + \epsilon F_2) = Q_{A_\epsilon}(I - S)^{1/2}$$

(17) $$= Q_{A_\epsilon}\left(I - \tfrac{1}{2}S - \tfrac{1}{8}S^2 - \cdots - \binom{\frac{1}{2}}{p}S^p - \cdots\right),$$

where the matrix series converges if $\rho(S) = \sigma_1(S) < 1$.

Also, asymptotically, when $\epsilon \to 0$, Q_{A_ϵ} is the unitary matrix with range $R(A_\epsilon)$ which is closest to Q_A.

From the well-known inequalities for singular values [15, p. 30],

$$\sigma_k(A + B) \leqq \sigma_k(A) + \sigma_k(B), \qquad \sigma_k(AB) \leqq \sigma_k(A)\sigma_k(B), \qquad k = 1, 2, \cdots,$$

we get

(18) $$\sigma_1(F) \leqq \sigma_1(E)\sigma_1(H_A^{-1}) \leqq \sigma_1(E)/(\sigma_p(A) - \epsilon\sigma_1(E_1)).$$

Since certainly $\sigma_1(E) \geqq \sigma_1(E_1)$, a sufficient condition for convergence of (17) is that

$$\epsilon\sigma_1(E)/\sigma_p(A) \leqq \tfrac{1}{2}.$$

Premultiplying (17) by P_B, we get

$$P_B Q_{A_\epsilon} = P_B Q_A + \epsilon P_B (I - P_A) F + P_B Q_{A_\epsilon}\left(\tfrac{1}{2}S + \tfrac{1}{8}S^2 + \cdots + \binom{\frac{1}{2}}{p}S^p + \cdots\right)$$

from which we derive the inequality

$$|\sigma_k(P_B Q_{A_\epsilon}) - \sigma_k(P_B Q_A)| \leqq \epsilon \cdot \sigma_1(P_B(I - P_A)) + \tau,$$

where

$$\tau = \tfrac{1}{2}\sigma_1(S) + \tfrac{1}{8}\sigma_1^2(S) + \cdots \leqq \tfrac{1}{2}\sigma_1(S)/(1 - \sigma_1(S)).$$

Now, we have $\sigma_1(P_B(I - P_A)) = \sin\theta_{\max}$ and, estimating $\sigma_1(S)$ and $\sigma_1(F)$ by (16) and (18), it follows that

$$(19) \qquad |\Delta\cos\theta_k| \leqq \delta \cdot \sin\theta_{\max} + O(\delta^2), \qquad \delta = \epsilon\sigma_1(E)/\sigma_p(A).$$

If instead we premultiply (17) by $(I - P_B)$ and proceed in the same way, we arrive at

$$(20) \qquad |\Delta\sin\theta_k| \leqq \delta \cdot \cos\theta_{\min} + O(\delta^2), \qquad \delta = \epsilon\sigma_1(E)/\sigma_p(A).$$

Now, assume that both A and B are perturbed by δA and δB, respectively, where

$$||\delta A||_2/||A||_2 \leqq \epsilon_A, \qquad ||\delta B||_2/||B||_2 \leqq \epsilon_B.$$

Then

$$\cos\theta_k = [\sigma_k(P_{B_\epsilon}Q_{A_\epsilon}) - \sigma_k(P_{B_\epsilon}Q_A)] + [\sigma_k(P_{B_\epsilon}Q_A) - \sigma_k(P_B Q_A)]$$

and, from (19), we get

$$|\Delta\cos\theta_k| \leqq \epsilon_A\kappa(A)\sin\theta_{\max}(A, B_\epsilon) + \epsilon_B\kappa(B)\sin\theta_{\max}(A, B) + O(\delta^2),$$

where

$$(21) \qquad \kappa(A) = \sigma_1(A)/\sigma_p(A), \qquad \kappa(B) = \sigma_1(B)/\sigma_q(B), \qquad \delta = \epsilon_A\kappa(A) + \epsilon_B\kappa(B).$$

A corresponding estimate holds for $|\Delta\sin\theta_k|$. Obviously, we have $\theta(A, B_\epsilon) = \theta(A, B) + O(\delta)$, and, thus, these estimates can be simplified to

$$\left|\Delta\begin{Bmatrix}\cos\theta_k \\ \sin\theta_k\end{Bmatrix}\right| \leqq \delta\begin{Bmatrix}\sin\theta_{\max} \\ \cos\theta_{\min}\end{Bmatrix} + O(\delta^2).$$

Combining these two estimates yields

$$|2 \cdot \sin\tfrac{1}{2}\Delta\theta_k| \leqq \delta g(\hat\theta_k) + O(\delta^2)$$

where

$$g(\theta) = \min\left(\frac{\sin\theta_{\max}}{\sin\theta}, \frac{\cos\theta_{\min}}{\cos\theta}\right), \qquad \hat\theta_k = \theta_k + \tfrac{1}{2}\Delta\theta_k.$$

The maximum of $g(\theta)$ for $0 \leqq \theta \leqq \pi/2$ is attained for

$$\theta = \arctan(\sin\theta_{\max}/\cos\theta_{\min}).$$

Since $2 \cdot \sin\tfrac{1}{2}\Delta\theta = \Delta\theta + O(\Delta\theta^3)$, it follows that

$$(22) \qquad |\Delta\theta_k| \leqq g_{\max}(\epsilon_A\kappa(A) + \epsilon_B\kappa(B)) + O(\delta^2),$$

where

$$g_{\max} = (\sin^2\theta_{\max} + \cos^2\theta_{\min})^{1/2} \leqq 2^{1/2}.$$

We conclude that when both $\kappa(A)$ and $\kappa(B)$ are small, then the angles θ_k are well determined.

We note that, if the columns in A are scaled, then $\kappa(A)$ will change but not $R(A)$. Also, the numerical algorithms for the QR-decomposition have the property that, unless column pivoting is used, they give the same *numerical* results independent of such a scaling. Therefore, it is often more relevant to take in (21) as condition number for A the number

$$\kappa'(A) = \min_D \kappa(AD), \qquad D = \mathrm{diag}(d_1, \cdots, d_p).$$

It has been shown in [20] and [21] that $\kappa(AD)$ is not more than a factor of $p^{1/2}$ away from its minimum, if in AD all columns have equal L_2-norm. This suggests that A and B should be assumed to be preconditioned so that

$$||a_i||_2 = ||b_j||_2 = 1, \qquad i = 1, \cdots, p, j = 1, \cdots, q.$$

We remark that $\kappa'(A)$ is essentially the spanning precision of the basis in $R(A)$ provided by A as defined in [21].

5. Numerical Methods. We assume in this section that the columns in A and B are linearly independent. The singular and near singular case will be briefly discussed in Section 6. For convenience, we also assume that A and B are real matrices, although all algorithms given here can easily be generalized to the complex case. Computed quantities will be marked by a bar.

In order to get the orthogonal bases for F and G, we need the QR-decompositions of the matrices A and B. We now describe two efficient methods for computing these. In the method of *Householder triangularizations* (HT) [16], orthogonal transformations of the type $Q_k = I - 2w_k w_k^T$ are used, where

$$w_k = (0, \cdots, 0, w_{kk}, \cdots, w_{mk})^T, \qquad ||w_k||_2 = 1.$$

The $m \times p$ matrix A is reduced to triangular form using premultiplications

$$Q_p \cdots Q_2 Q_1 A = \left(\frac{R_A}{0}\right) \begin{matrix} \} p \\ \} m - p \end{matrix}$$

where w_k is chosen so that Q_k annihilates the appropriate elements in the kth column. Since $Q_k^{-1} = Q_k$, orthogonal bases Q_A for $R(A)$ can then be computed by premultiplying the first p columns in the unit matrix I_m by the same transformations in reversed order,

$$Q_A = Q_1 Q_2 \cdots Q_p \left(\frac{I_p}{0}\right).$$

For this method, a very satisfactory error analysis is given in [23].

Assume that floating point arithmetic with a mantissa of t binary digits is used, and that inner-products are accumulated in double precision wherever possible. Then, there exists an *exactly* orthogonal matrix Q such that the computed matrices satisfy

$$(23) \qquad Q^T(A + E_A) = \left(\frac{\bar{R}_A}{0}\right), \qquad \bar{Q}_A = Q\left(\frac{I_p}{0}\right) + F_A = Q_{\bar{A}} + F_A,$$

$$||E_A||_F = 12.5p2^{-t} ||A||_F, \qquad ||F_A||_F = 12.5p^{3/2}2^{-t},$$

where $Q_{\bar{A}}$ is an exactly orthogonal basis for $R(A + E_A)$. From this and a similar estimate for $\bar{Q}_B$, we get

(24) $\qquad |\sigma_k(\bar{M}) - \sigma_k(\tilde{M})| \leq \sigma_1(\bar{M} - \tilde{M}) \leq 13.0(p^{3/2} + q^{3/2})2^{-t}$,

where $\tilde{M} = \bar{Q}_A^T \bar{Q}_{\bar{B}}$ and the constant 13.0 accounts for the rounding errors in computing the product $\bar{Q}_A^T \bar{Q}_B$. We have $\sigma_k(\tilde{M}) = \cos \tilde{\theta}_k$, where $\tilde{\theta}_k$ are the exact angles between $(A + E_A)$ and $(B + E_B)$. Thus, the difference between $\tilde{\theta}_k$ and θ_k can be estimated from (22),

(25) $\qquad |\tilde{\theta}_k - \theta_k| \leq 12.5 \cdot 2^{1/2} \cdot \delta + O(\delta^2), \qquad \delta = (p\kappa(A) + q\kappa(B))2^{-t}$.

Finally, the errors $\bar{\sigma}_k(\bar{M}) - \sigma_k(\bar{M})$ in computing the singular values of $\bar{M}$, using the procedure in [17], will be of the same order of magnitude as those in (24).

The error estimate given above is satisfactory, except when $\theta_k \ll 1$. In this case, the errors in $\cos \theta_k$ from (24) will give rise to errors in θ_k which may be much larger than those in (25). We return later to the problem of accurately computing small angles.

An orthogonal basis Q_A' for $[R(A)]^- = N(A^T)$ can be obtained by applying the transformations Q_k, $k = p, \cdots, 1$, to the last $(m - p)$ columns in I_m,

$$Q_A' = Q_1 Q_2 \cdots Q_p \left(\frac{0}{I_{m-p}}\right).$$

Also, in this case, the estimate (23) for $\bar{Q}_A'$, (24) and (25) still hold if the factor $p^{3/2}$ is everywhere replaced by $p(m - p)^{1/2}$.

The QR-decomposition of a matrix A can also be computed using the modified Gram-Schmidt method (MGS) [2]. The matrix A is then transformed in p steps, $A = A_1, A_2, \cdots, A_{p+1} = Q_A$ where

$$A_k = (q_1, \cdots, q_{k-1}, a_k^{(k)}, \cdots, a_p^{(k)}).$$

The matrix A_{k+1}, $k = 1, 2, \cdots, p$, is computed by

$$q_k = a_k^{(k)}/||a_k^{(k)}||_2, \qquad a_j^{(k+1)} = (I - q_k q_k^T)a_j^{(k)}, \qquad j > k,$$

and the elements in the kth row of R_A are

$$r_{kk} = ||a_k^{(k)}||_2, \qquad r_{kj} = q_k^T a_j^{(k)}, \qquad j > k.$$

It has been shown in [2, pp. 10, 15] that the computed matrices $\bar{R}_A$ and $\bar{Q}_A$ satisfy

(26)
$$A + E_A = \bar{Q}_A \bar{R}_A, \qquad ||E_A||_F \leq 1.5(p - 1)2^{-t} ||A||_F,$$
$$||Q_{\bar{A}} - \bar{Q}_A||_2 \leq p(p + 1)\kappa(A) \cdot 2^{-t},$$

where $Q_{\bar{A}}$ is an exactly orthogonal basis for $R(A + E_A)$ and quantities of order $\kappa^2(A)2^{-2t}$ have been neglected. With MGS, $\bar{Q}_A$ will in general not be orthogonal to working accuracy, and, therefore, we cannot hope to get principal vectors which are nearly orthogonal. Also, the condition numbers $\kappa(A)$ and $\kappa(B)$ will enter in the estimate corresponding to (24). However, since $\kappa(A)$ and $\kappa(B)$ already appear in (25), we can hope to get the principal angles as accurately as with HT. Experimental results reported in Section 7 indicate that this actually is the case.

An advantage with MGS is that the total number of multiplications required to compute $\bar{R}_A$ and $\bar{Q}_A$ is less than for HT, i.e.,

$$\text{MGS: } p^2m, \qquad \text{HT: } 2p^2(m - p/3).$$

If only the principal angles are wanted, then the number of multiplications in the SVD-algorithm is approximately $2q^2(p - q/3)$. Thus, when $m \gg p$, the dominating work is in computing Q_A and Q_B and, in this case, MGS requires only half as much work as HT. If also the principal vectors are wanted, we must compute the full SVD of $M = Y_A C y_B^H$. Assuming two iterations per singular value, this requires approximately $7q^2(p + 10q/21)$ multiplications. To compute $U_A = Q_A Y_A$ and $U_B = Q_B Y_B$ a further $mq(p + q)$ multiplications are needed.

To get a basis for $[R(A)]^-$ using MGS, we have to apply the method to the bordered matrix $(A \mid I_m)$, and, after m steps, pick out $(m - p)$ appropriate columns. Especially when $(m - p) \ll m$, the number of multiplications compares unfavourably with HT,

$$\text{MGS: } m^2(m + 2p), \qquad \text{HT: } 2mp(m - p) + \tfrac{2}{3}p^3.$$

In some applications, e.g. canonical correlations, we want to express the principal vectors as linear combinations of the columns in A and B, respectively. We have $U_A = Q_A Y_A = A(R_A^{-1} Y_A)$, and hence

$$U_A = A X_A, \qquad U_B = B X_B,$$

where

(27) $$X_A = R_A^{-1} Y_A, \qquad X_B = R_B^{-1} Y_B.$$

We remark that if we let $\bar{X}_A$ and $\bar{X}_B$ denote the computed matrices, then $A\bar{X}_A$ and $B\bar{X}_B$ will *not* in general be orthogonal to working accuracy even when HT is used.

We now turn to the problem of accurately determining small angles. One method is to compute $\sin \theta_k$ from the SVD (13) of the matrix

(28) $$G = (I - P_A)Q_B = Q_B - Q_A M.$$

If we put $\tilde{G} = Q_{\bar{B}} - Q_{\bar{A}}\bar{M}$, then, using $\bar{Q}_A = Q_{\bar{A}} + F_A$, we get

$$\bar{Q}_B + \bar{Q}_A(\bar{Q}_A^T Q_B) = \tilde{G} + (I - Q_{\bar{A}}Q_{\bar{A}}^T)F_B + (Q_{\bar{A}}F_A^T + F_A Q_{\bar{A}}^T)(Q_{\bar{B}} + F_B).$$

Neglecting second order quantities,

$$\|\bar{Q} - \tilde{G}\|_2 \leq \|F_B\|_2 + 2\|F_A\|_2 + 2q^{1/2}2^{-t},$$

where the last term accounts for the final rounding of the elements in $\bar{M}$ and $\bar{G}$. Thus, if $\bar{Q}_A$ and $\bar{Q}_B$ are computed by HT, we have, from (23),

$$|\sigma_k(\bar{G}) - \sigma_k(\tilde{G})| \leq 13.2(q^{3/2} + 2p^{3/2})2^{-t}.$$

It follows that the singular values of the computed matrix $\bar{G}$ will differ little from $\sin \bar{\theta}_k$, and, thus, small angles will be as accurately determined as is allowed by (25). From (26), the corresponding error estimate for MGS is obtained. In the spirit of the modified Gram-Schmidt method, the matrix G should be computed as

(29) $$G = (I - q_1 q_1^T) \cdots (I - q_p q_p^T)Q_B.$$

Computational experience indicates that this gives much improved accuracy in $\sin \theta$ when $\kappa(A) \gg \kappa(B)$.

Since the matrix G is $m \times q$, computing the singular values of G when $m \gg q$ will require about $2mq^2$ multiplications. If Y_A has been computed, then the SVD of G may be computed with only mq^2 multiplications from

(30) $$G Y_B = (I - P_A)Q_B Y_B = W_A S.$$

Moreover, if U_A and U_B are available, we can obtain $\sin \theta$ from

(31) $$(U_B - U_A C)^T (U_B - U_A C) = S^2$$

or, alternatively,

(32) $$(U_B - U_A)^T (U_B - U_A) = 2(I - C).$$

From the last equation, we can compute $2 \sin \frac{1}{2}\theta_k = (2(1 - \cos \theta_k))^{1/2}$, which, since $0 \leq \frac{1}{2}\theta_k \leq \pi/4$, accurately determines both $\sin \theta_k$ and $\cos \theta_k$.

We finally remark about an apparent imperfection of MGS. When $A = B$ (exactly), we will obviously get $\bar{Q}_A = \bar{Q}_B$. The exact angle equals zero, and HT will always give computed angles near zero. This is not true for MGS, however, since we only have the estimate

$$||I - \bar{Q}_A^T \bar{Q}_A||_2 \leq 2p(p + 1)\kappa(A)2^{-t}.$$

Therefore, the singular values of $\bar{M} = \bar{Q}_A^T \bar{Q}_A$ may not be near one when $\kappa(A)$ is large. If, however, only $A \simeq B$, then the rounding errors in computing Q_A and Q_B will not be correlated, and in an ill-conditioned case, we will probably not get all angles near zero either with HT or MGS.

When $A = B$, then $\bar{M} = \bar{Q}_A^T \bar{Q}_A$ will be symmetric and, thus, SVD will give $\bar{Y}_A \simeq \bar{Y}_B$ and, therefore, $\bar{U}_A \simeq \bar{U}_B$ also with MGS. It follows that, if (32) is used, MGS will always yield angles near zero in this case.

We have not tried to determine error estimates for the methods based on (30)–(32). On the test matrices described in Section 7, the method based on (28) gave significantly more accurate results, especially for the more well-conditioned angles.

6. The Singular Case. We now consider the case when A and/or B does not have full column rank. In this case, the problem of computing principal angles and vectors is not well posed, since arbitrarily small perturbations in A and B will change the rank of A and/or B. The main computational difficulty then lies in assigning the correct rank to A and B. The most satisfactory way of doing this generally is the following [10]. Let the $m \times p$ matrix A have the SVD

$$A = Q_A D_A V_A^T, \qquad D_A = \text{diag}(\sigma_k(A)).$$

Let ϵ be a suitable tolerance and determine $p' \leq p$ from

(33) $$\sum_{i = p'+1}^{n} \sigma_i^2(A) \leq \epsilon^2 < \sum_{i=p'}^{n} \sigma_i^2(A).$$

We then approximate A with an $m \times p$ matrix A' such that $\text{rank}(A') = p'$,

$$A' = (Q_A' Q_A'')\begin{pmatrix} D_A' & 0 \\ 0 & 0 \end{pmatrix}(V_A' V_A'')^T, \qquad D_A' = \text{diag}(\sigma_1, \cdots, \sigma_{p'}),$$

where $Q_A = (Q'_A Q''_A)$, $V_A = (V'_A V''_A)$ have been partitioned consistently with the diagonal matrix. The matrix B is approximated in the same way.

If, instead of (1), we assume that

$$p' = \operatorname{rank}(A') \geqq \operatorname{rank}(B') = q' \geqq 1,$$

then we can compute the principal angles and vectors associated with $R(A')$ and $R(B')$ by the previously derived algorithms, where now Q'_A and Q'_B should replace Q_A and Q_B.

In order to express the principal vectors of $R(A')$ as linear combinations of columns in A', we must solve the compatible system

$$A' X_A = U_A = Q'_A Y_A.$$

Since V''_A is an orthogonal basis for $N(A)$, the general solution can be written

$$X_A = V'_A D'^{-1}_A Y_A + V''_A C_A,$$

where C_A is an arbitrary matrix. It follows that, by taking $C_A = 0$, we get the unique solution which minimizes $\|X_A\|_F$; cf. [17]. Thus, we should take

$$(34) \qquad X_A = V'_A D'^{-1}_A Y_A, \qquad X_B = V'_B D'^{-1}_B Y_B,$$

where X_A is $p \times p'$ and X_B is $q \times q'$.

The approach taken above also has the advantage that only one decomposition, the SVD, is used throughout. It can, of course, also be used in the nonsingular case. However, computing the SVD of A and B requires much more work than computing the corresponding QR-decompositions. In order to make the QR-methods work also in the singular case, column pivoting must be used. This is usually done in such a way ([2], [12] and [16]) that the triangular matrix $R = (r_{ij})$ satisfies

$$|r_{kk}|^2 \geqq \sum_{i=k}^{i} |r_{ij}|^2, \qquad k < j \leqq n.$$

Such a triangular matrix is called normalized, and, in particular, the sequence $|r_{11}|$, $|r_{22}|, \cdots, |r_{pp}|$ is nonincreasing. In practice, it is often satisfactory to take the numerical rank of A to be p' if for a suitable tolerance ϵ we have

$$(35) \qquad |r_{p'p'}| > \epsilon \geqq |r_{p'+1,p'+1}|.$$

We then approximate $A = Q_A R_A$ by a matrix $A' = Q_A R'_A$ of rank p' by putting

$$r'_{ij} = r_{ij}, \qquad i \leqq p', \qquad r'_{ij} = 0, \qquad i > p'.$$

It has been shown in [14] how to obtain the solution (32) of minimum length from this decomposition.

If we use the criterion (33), there is a risk of choosing p' too large. Indeed, it seems difficult to improve on the inequalities [12]

$$(36) \qquad 3(4^k + 6k - 1)^{-1/2} |r_{kk}| \leqq \sigma_k(A) \leqq (n + k + 1)^{1/2} |r_{kk}|$$

from which it is seen that $\sigma_k(A)$ may be smaller than $|r_{kk}|$ by a factor of magnitude 2^{-k}.

However, this rarely occurs in practice. Often the inequality

$$\kappa(A) \geqq |r_{11}| / |r_{pp}|,$$

represents a pretty good approximation to the condition number for the nonsingular case.

7. Test Results. The algorithms in Section 5 have been tested on the UNIVAC 1108 of Lund University. Single precision floating-point numbers are represented by a normalized 27 bit mantissa, whence the machine precision is equal to $2^{-26} \approx 1.5 \cdot 10^{-8}$.

For the tests, we have taken $F = R(A)$, where A is the $m \times p$ matrix

$$A = \frac{1}{k^{1/2}} \begin{bmatrix} e & 0 & \cdots & 0 \\ 0 & e & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & e \end{bmatrix}, \qquad e = \left. \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} \right\} m/p$$

and m/p is an integer. Thus, A is already orthogonal, and we can take $Q_A = A$. Further, $G = R(B)$ where B is the $m \times p$ Vandermonde matrix

$$B = \begin{bmatrix} 1 & x_0 & \cdots & x_0^{p-1} \\ 1 & x_1 & \cdots & x_1^{p-1} \\ & & \cdots & \\ 1 & x_{m-1} & \cdots & x_{m-1}^{p-1} \end{bmatrix}, \qquad x_i = -1 + 2i/(m+1).$$

The condition number $\kappa(B)$ is known to grow exponentially with p, when the ratio m/p is kept constant. These matrices A and B are the ones appearing in [6]. There is exactly one vector, $u = (1, 1, \cdots, 1)^T$, which belongs to both F and G, so there will be one minimum angle $\theta = 0$.

For the tests, the matrix B was generated in double precision and then rounded to single precision. The procedures used for the QR-decompositions are apart from minor details identical with procedures published in [3] and [5]. The columns were implicitly scaled to have equal L_2-norm and column pivoting was performed. Inner products were *not* accumulated in double precision. For checking purposes, a three term recurrence relation [8] was used in double precision to compute an exact single precision orthogonal basis for $R(B)$.

For $m/p = 2$ and $p = 5(2)17$, Q_A was computed both by the method of Householder and the modified Gram-Schmidt method. Then $\cos \theta_k$, Y_A and Y_B were com-

TABLE 1

		Householder			Gram-Schmidt		
m	p	$F(\bar{U}_A)$	$F(\bar{U}_B)$	$m(\cos \bar{\theta}_k)$	$F(\bar{U}_A)$	$F(\bar{U}_B)$	$m(\cos \bar{\theta}_k)$
10	5	15	21	5	20	18	4
14	7	22	35	8	27	33	10
18	9	40	27	10	37	437	41
22	11	42	29	38	40	1130	49
26	13	58	48	1416	47	9013	621
30	15	62	51	2535	60	55322	1758
34	17	68	60	7582	76	788466	32650

TABLE 2

$m = 26$ k	$p = 13$ $\cos \theta_k$	Householder $\Delta \cos \bar{\theta}_k \cdot 10^8$	Gram-Schmidt $\Delta \cos \bar{\theta}_k \cdot 10^8$
1	0.99999991	2	2
2	0.99823275	0	51
3	0.99814397	$-$ 32	-135
4	0.99032703	6	-137
5	0.98988846	5	351
6	0.97646081	0	$-$ 58
7	0.96284604	38	21
8	0.94148906	$-$ 5	$-$ 10
9	0.91758607	$-$ 31	$-$ 40
10	0.87013717	25	-290
11	0.76365752	1416	620
12	0.06078817	106	$-$ 18
13	0.01558526	$-$ 52	$-$ 55

TABLE 3

		Householder[1]			Gram-Schmidt[1] from (29)	from (28)
m	p	$m(\sin \bar{\theta}_k)$	$m(\sin \tilde{\theta}_k)$	$m(\sin \bar{\theta}_k)$	$m(\sin \bar{\theta}_k)$	$m(\sin \tilde{\theta}_k)$
10	5	3	2	3	3	6
14	7	7	8	3	4	8
18	9	32	31	55	9	87
22	11	141	142	46	39	612
26	13	1661	1662	366	517	5902
30	15	2919	2912	1290	1355	32537
34	17	7604	7608	37284	798	126731

[1] $\sin \bar{\theta}_k$ computed as $\sigma_k((I - P_A)Q_B)$, $\sin \tilde{\theta}_k$ as $\sigma_k((I - P_B)Q_A)$.

puted by the procedure in [17], and finally U_A and U_B from (11). The results are shown in Table 1, where

$$m(\sigma_k) = \max_k |\sigma_k - \bar{\sigma}_k| \cdot 10^8, \qquad F(U) = ||I - U^T U||_F \cdot 10^8.$$

Notice, that because of rounding errors in the computation of the SVD, the values σ_k are not exact to single precision.

For the Gram-Schmidt method, the predicted lack of orthogonality in U_B, when $\kappa(B)$ is large, is evident. However, there is no significant difference in the accuracy of $\cos \bar{\theta}_k$ between the two methods. In Table 2, we show for $m = 26$ and $p = 13$ the errors in $\cos \bar{\theta}_k$ for each k.

For the same values of m and p, $\sin \theta_k$ were computed from the singular values of both the matrix $(I - P_A)Q_B$ and the matrix $(I - P_B)Q_A$. The results in Table 3 again show no significant difference in accuracy between the two methods. For the Gram-Schmidt method, the values of $\sin \theta_k$ differ somewhat between the two matrices,

TABLE 4

$m = 26$	$p = 13$	Householder[1]		Gram-Schmidt[1]	
k	$\sin \theta_k$	$\Delta \sin \bar{\theta}_k \cdot 10^8$	$\Delta \sin \tilde{\theta}_k \cdot 10^8$	$\Delta \sin \bar{\theta}_k \cdot 10^8$	$\Delta \sin \tilde{\theta}_k \cdot 10^8$
1	0.00000000	2	5	2	4
2	0.05942260	0	0	0	− 2
3	0.06089681	485	484	69	76
4	0.13875174	56	55	2	22
5	0.14184706	− 30	− 29	− 23	− 71
6	0.21569431	32	31	− 35	− 20
7	0.27005038	− 127	− 127	− 20	− 26
8	0.33704301	26	25	− 2	3
9	0.39753669	91	90	− 25	− 5
10	0.49280931	− 44	− 42	260	318
11	0.64562106	−1661	−1662	−365	−517
12	0.99815036	5	11	−158	15
13	0.99987821	14	20	90	13

[1] $\sin \bar{\theta}_k$ computed as $\sigma_k((I - P_A)Q_B)$, $\sin \tilde{\theta}_k$ as $\sigma_k((I - P_B)Q_A)$.

whereas the corresponding values for the Householder method are almost identical· This is confirmed by Table 4, where, again for $m = 26$, $p = 13$, results for each k are shown. For the Gram-Schmidt method the matrix $(I - P_B)Q_A$ was computed both from (28) as $Q_A - Q_B(Q_B^T Q_A)$ and from (29) as $\prod (I - q_k q_k^T)Q_B$. The results in Table 3 clearly show that (29) should be used.

The authors are very pleased to acknowledge the help of Mr. Jan Svensson, who carried out most of the tests described in this section.

Department of Mathematics
Linköping University
S-581 83 Linköping, Sweden

Computer Science Department
Stanford University
Stanford, California 94305

1. S. N. Afriat, "Orthogonal and oblique projectors and the characteristics of pairs of vector spaces," *Proc. Cambridge Philos. Soc.*, v.53, 1957, pp. 800–816. MR **20** #1389.
2. Å. Björck, "Solving linear least squares problems by Gram-Schmidt orthogonalization," *Nordisk Tidskr. Informationsbehandling (BIT)*, v. 7, 1967, pp. 1–21. MR **53** #5126.
3. Å. Björck, "Iterative refinement of linear least squares solutions. II," *Nordisk Tidskr. Informationsbehandling (BIT)*, v. 8, 1968, pp. 8–30. MR **39** #1094.
4. Å. Björck & C. Bowie, "An iterative algorithm for computing the best estimate of an orthogonal matrix," *SIAM J. Numer. Anal.*, v. 8, 1971, pp. 358–364.
5. P. Businger & G. H. Golub, "Handbook series linear algebra. Linear least squares solutions by Householder transformations," *Numer. Math.*, v. 7, 1965, pp. 269–276. MR **31** #862.
6. C. Cohen & A. Ben-Israel, "On the computation of canonical correlations," *Cahiers Centre Études Recherche Opér.*, v. 11, 1969, pp. 121–132. MR **40** #6683.
7. C. Cohen, *An Investigation of the Geometry of Subspaces for Some Multivariate Statistical Models*, Thesis, Dept. of Indust. Eng., University of Illinois, Urbana, Ill., 1969.

8. G. DAHLQUIST, B. SJÖBERG & P. SVENSSON, "Comparison of the method of averages with the method of least squares," *Math. Comp.*, v. 22, 1968, pp. 833–845. MR **39** #1099.

9. C. DAVIES & W. KAHAN, "The rotation of eigenvectors by a perturbation. III," *SIAM J. Numer. Anal.*, v. 7, 1970, pp. 1–46.

10. C. ECKART & G. YOUNG, "The approximation of one matrix by another of lower rank," *Psychometrika*, v. 1, 1936, pp. 211–218.

11. C. ECKART & G. YOUNG, "A principal axis transformation for non-Hermitian matrices," *Bull. Amer. Math. Soc.*, v. 45, 1939, pp. 118–121.

12. D. K. FADDEEV, V. N. KUBLANOVSKAJA & V. N. FADDEEVA, *Sur les Systèmes Linéaires Algébriques de Matrices Rectangulaires et Mal-Conditionnées* (Actes Colloq. Internat. C.N.R.S. No. 165, Besançon, 1966) Éditions C.N.R.S., Paris, 1968, pp. 161–170. MR **37** #6017.

13. G. FIX & R. HEIBERGER, "An algorithm for the ill conditioned generalized eigenvalue problem," *Numer. Math.* (To appear.)

14. R. J. HANSON & C. L. LAWSON, "Extensions and applications of the Householder algorithm for solving linear squares problems," *Math. Comp.*, v. 23, 1969, pp. 787–812. MR **41** #2905.

15. I. C. GOHBERG & M. G. KREĬN, *Introduction to the Theory of Linear Nonselfadjoint Operators*, "Nauka," Moscow, 1965; English transl. Transl. Math. Monographs, vol. 18, Amer. Math. Soc., Providence, R. I., 1969. MR **39** #7447.

16. G. H. GOLUB, "Numerical methods for solving linear least squares problems," *Numer. Math.*, v. 7, 1965, pp. 206–216. MR **31** #5323.

17. G. H. GOLUB & C. REINSCH, "Singular value decompositions and least squares solutions," *Numer. Math.*, v. 14, 1970, pp. 403–420.

18. I. J. GOOD, *The Estimation of Probabilities. An Essay on Modern Bayesian Methods*, Research Monograph No. 30, M.I.T. Press, Cambridge, Mass., 1965, pp. 87–89. MR **32** #3186.

19. H. HOTELLING, "Relations between two sets of variates," *Biometrika*, v. 28, 1936, pp. 321–377.

20. A. VAN DER SLUIS, "Condition numbers and equilibration of matrices," *Numer. Math.*, v. 14, 1969/1970, pp. 14–23. MR **40** #6760.

21. J. M. VARAH, "Computing invariant subspaces of a general matrix when the eigensystem is poorly conditioned," *Math. Comp.*, v. 24, 1970, pp. 137–149. MR **41** #9434.

22. P. Å WEDIN, *On Pseudoinverses of Perturbed Matrices*, Dept. of Comput. Sci., Lund University, Lund, Sweden, 1969.

23. J. WILKINSON, "Error analysis of transformations based on the use of matrices of the form $I - 2ww^H$," *Error in Digital Computation*, vol. II (Proc. Sympos. Math. Res. Center, U.S. Army, Univ. of Wisconsin, Madison, Wis., 1965), Wiley, New York, 1965, pp. 77–101. MR **32** #6711.