SIAM J. NUMER. ANAL.
Vol. 13, No. 1, March 1976

GENERALIZING THE SINGULAR VALUE DECOMPOSITION*

CHARLES F. VAN LOAN†

Abstract. Two generalizations of the singular value decomposition are given. These generalizations provided a unified way of regarding certain matrix problems and the numerical techniques which are used to solve them.

1. Introduction and notation.

A matrix factorization having great importance in numerical linear algebra is the singular value decomposition:

THEOREM 1 (the singular value decomposition (SVD)). *If A is a real $m \times n$ matrix, then there exist orthogonal matrices $U(m \times m)$ and $V(n \times n)$ such that*

$$U^T A V = \operatorname{diag}(\mu_1, \cdots, \mu_q), \qquad q = \min\{m, n\},$$

where $\mu_1 \geqq \mu_2 \geqq \cdots \geqq \mu_r > \mu_{r+1} = \cdots = \mu_q = 0$ and $r = \operatorname{rank}(A)$.

The μ_i are called the singular values of A, and we define the set $\mu(A)$ by $\mu(A) = \{\mu_1, \cdots, \mu_q\}$. A proof of the SVD can be found in the book by Lawson and Hanson [7] together with a description of the standard algorithm used for computing it. This algorithm is due to Golub and others [4], [5], [6].

Of the many properties that the singular values possess, the following are of particular interest:

$$\mu \in \mu(A) \Rightarrow \det(A^T A - \mu^2 I_n) = 0, \tag{1.1}$$

$$\mu \in \mu(A) \Rightarrow \mu \text{ is a stationary value of } \frac{\|Ax\|}{\|x\|}. \tag{1.2}$$

(A summary of notation is given below.) A natural generalization of the singular value concept stems from each of these implications:

$$\text{Find those } \mu \geqq 0 \text{ for which } \det(A^T A - \mu^2 B^T B) = 0. \tag{1.1'}$$

$$\text{Find the stationary values of } \frac{\|Ax\|_S}{\|x\|_T}. \tag{1.2'}$$

Here the matrices A and B have an equal number of columns and the norms $\|\cdot\|_S$ and $\|\cdot\|_T$ are defined below.

The purpose of this paper is to discuss (1.1′) and (1.2′) and the generalized singular value decompositions that they each suggest. By doing this we hope to provide a new way of regarding certain problems and the numerical techniques used to solve them. We feel that a sharp delineation of the generalized singular

* Received by the editors October 4, 1974, and in revised form May 13, 1975.

† Department of Mathematics, University of Manchester, Manchester, England M13 9PL. This work was supported in part by a grant from the Science Research Council.

value concept is called for even though the idea has been around, albeit implicitly, for a number of years. Our notation is as follows:

$R^{m\times n}$ = vector space of real $m\times n$ matrices,

$R^n = R^{n\times 1}$ = vector space of real n-tuples,

$I_n = n\times n$ identity matrix,

A^T = "A transpose" $(A=(a_{ij})\in R^{m\times n} \Rightarrow A^T=(a_{ji})\in R^{n\times m})$,

$NS(A)=\{x|Ax=0\}$ = nullspace of A,

$\|x\|=(x^Tx)^{1/2}$, $x\in R^n$,

$\|x\|_P=(x^TPx)^{1/2}$, $x\in R^n$, $P\in R^{n\times n}$ positive definite.

Finally, $D=(d_{ij})\in R^{m\times n}$ is diagonal whenever $d_{ij}=0$ for all $i\neq j$. When this is the case, we may write $D=\operatorname{diag}(\mu_1,\cdots,\mu_q)$ where $q=\min\{m,n\}$ and $\mu_i=d_{ii}$ for $i=1,\cdots,q$.

2. Two generalized singular value decompositions. In this section we produce two matrix decompositions which are relevant to the problems (1.1′) and (1.2′). Let us begin by defining the concept of "B-singular values".

DEFINITION 1. The *B-singular values* of a matrix A are elements of the set $\mu(A,B)$ defined by

$$\mu(A,B)=\{\mu\,|\,\mu\geqq 0, \det(A^TA-\mu^2B^TB)=0\},$$

where $A\in R^{m_a\times n}$, $B\in R^{m_b\times n}$ and $m_a\geqq n$. (As we shall see, the assumption $m_a\geqq n$ is not restrictive from the applications point of view.)

The following decomposition is a generalization of the SVD, and from it one can easily obtain the B-singular values.

THEOREM 2 (the B-singular value decomposition (BSVD). *Suppose $A\in R^{m_a\times n}$, $B\in R^{m_b\times n}$ and $m_a\geqq n$. There exist orthogonal matrices $U(m_a\times m_a)$ and $V(m_b\times m_b)$ and a nonsingular matrix $X(n\times n)$ such that*

$$U^TAX=D_A=\operatorname{diag}(\alpha_1,\cdots,\alpha_n),\qquad \alpha_i\geqq 0,$$
$$V^TBX=D_B=\operatorname{diag}(\beta_1,\cdots,\beta_q),\qquad \beta_i\geqq 0, \tag{2.1}$$

where $q=\min\{m_b,n\}$, $r=\operatorname{rank}(B)$ and $\beta_1\geqq\cdots\geqq\beta_r>\beta_{r+1}=\cdots=\beta_q=0$. If $\alpha_j=0$ for any j, $r+1\leqq j\leqq n$, then $\mu(A,B)=\{\mu\,|\,\mu\geqq 0\}$. Otherwise, $\mu(A,B)=\{\alpha_i/\beta_i\,|\,i=1,\cdots,r\}$.

Proof. Let the SVD of $\begin{pmatrix}A\\B\end{pmatrix}$ be given by

$$Q^T\begin{pmatrix}A\\B\end{pmatrix}Z_1=\operatorname{diag}(\gamma_1,\cdots,\gamma_n), \tag{2.2}$$

where $\gamma_1\geqq\cdots\geqq\gamma_k>\gamma_{k+1}=\cdots=\gamma_n=0$, $k=\operatorname{rank}\begin{pmatrix}A\\B\end{pmatrix}$. By defining $D=\operatorname{diag}(\gamma_1,\cdots,\gamma_k)\in R^{k\times k}$ and partitioning the orthogonal $Z_1\in R^{n\times n}$ as

$$Z_1=[\underbrace{Z_{11}}_{k}\,|\,\underbrace{Z_{12}}_{n-k}],$$

we find that $A_1^TA_1+B_1^TB_1=I_k$, where $A_1=AZ_{11}D^{-1}\in R^{m_a\times k}$ and $B_1=BZ_{11}D^{-1}\in R^{m_b\times k}$. If $r=\text{rank}\,(B)$ and

$$V^TB_1Z_2=\text{diag}\,(\beta_1,\cdots,\beta_p),\qquad p=\min\{m_b,k\},\tag{2.3}$$

denotes the SVD of B_1 with $\beta_1\geqq\cdots\geqq\beta_r>\beta_{r+1}=\cdots=\beta_p=0$, then

$$V^TBZ_1\begin{pmatrix}D^{-1}Z_2 & 0\\ 0 & I_{n-k}\end{pmatrix}=D_B\equiv\text{diag}\,(\beta_1,\cdots,\beta_q),\tag{2.4}$$

where $\beta_{p+1}=\cdots=\beta_q=0$ and $q=\min\{m_b,n\}$.

Next, we observe that the columns of A_1Z_2 are mutually orthogonal because

$$(A_1Z_2)^T(A_1Z_2)=Z_2^T(I_k-B_1^TB_1)Z_2=\text{diag}\,(1-\beta_1^2,\cdots,1-\beta_k^2).$$

Hence, there exists an orthogonal matrix $U\in R^{m_b\times m_b}$ such that

$$A_1Z_2=U\,\text{diag}\,(\alpha_1,\cdots,\alpha_k)\in R^{m_a x k}.\tag{2.5}$$

(For example, if $U^T(A_1Z_2)=R$ is the Householder upper triangularization of A_1Z_2, then it can be shown that the triangular matrix R is in fact diagonal.) We may assume that the α_i are nonnegative in (2.5). Defining $\alpha_i=0$ for $i=k+1,\cdots,n$ we see that

$$U^TAZ_1\begin{pmatrix}D^{-1}Z_2 & 0\\ 0 & I_{n-k}\end{pmatrix}=D_A=\text{diag}\,(\alpha_1,\cdots,\alpha_n).\tag{2.6}$$

Thus, (2.1) is established by setting $\begin{pmatrix}D^{-1}Z_2 & 0\\ 0 & I_{n-k}\end{pmatrix}=X$ in (2.4) and (2.6).

To show that the theorem correctly specifies $\mu(A,B)$, observe that

$$\det(A^TA-\mu^2B^TB)=\det(X)^{-2}\prod_{i=1}^{r}(\alpha_i^2-\mu^2\beta_i^2)\prod_{i=r+1}^{n}\alpha_i^2.\qquad\text{Q.E.D.}$$

Remark 1. *If $m_a<n$ and $m_b<n$*, then the diagonalization (2.1) may not exist as consideration of the example $A=(1\quad 0)$, $B=(0\quad 1)$ shows. Fortunately, this is not disastrous because $m_a\geqq n$ in all the examples to be discussed.

Remark 2. The transformation matrices U, V and X are not unique. In particular, we may replace X by XD in (2.1) whenever $D=\text{diag}\,(\delta_1,\cdots,\delta_n)\in R^{n\times n}$ with $\delta_i>0$. If A^TA and B^TB commute, then it is possible to choose X orthogonal. Relatedly, if $B=I_n$, then from (2.1), $XD_B^{-1}=V$ is orthogonal and $U^TA(XD_B^{-1})=D_AD_B^{-1}$ represents the SVD of A.

Remark 3. $\mu(A,B)=\{\mu\,|\,\mu\geqq 0\}\Leftrightarrow NS(A)\cap NS(B)\neq\{0\}$. In this case the subspace spanned by the first k columns of X is orthogonal to the subspace spanned by the last $n-k$ columns of X $\left(k=\text{rank}\begin{pmatrix}A\\B\end{pmatrix}\right)$. This latter subspace equals $NS(A)\cap NS(B)=NS\begin{pmatrix}A\\B\end{pmatrix}$.

Remark 4. If x_i denotes the ith column of X, then from (2.1) we have $\|Ax_i\|=\alpha_i$ $(i=1,\cdots,n)$ and $\|Bx_i\|=\beta_i$ $(i=1,\cdots,q)$. If the columns of X are scaled so as to have unit length, then this implies that each α_i is bounded below (above) by the smallest (largest) singular value of A. Similarly, each β_i is bounded below (above) by the smallest (largest) singular value of B.

The problem of finding the roots of $\det(A^TA-\mu^2B^TB)=0$ is clearly a synthesis of the problems $\det(A^TA-\mu^2I_n)=0$ and $\det(A-\lambda B)=0$. Indeed, the B-singular value problem inherits the difficulties associated with each of these more basic problems:

(a) When B is close to rank deficiency, it is difficult to compute the "stable" elements of $\mu(A, B)$ (that is, those B-singular values which are not wildly affected by small changes in B). This parallels the situation in the $A-\lambda B$ problem [9], [10].

(b) Formation of the products A^TA and B^TB can lead to a nontrivial loss of accuracy. This explains why one does not work with (A^TA) in the singular value problem.

A routine which overcomes these difficulties is the VZ algorithm [11]. This routine can find the stable elements of $\mu(A, B)$ without forming A^TA and B^TB, and its success is unaffected by rank degeneracy in B. On the negative side, VZ has no regard for symmetry and the transformation matrices U and V in (2.1) are ill-determined by the process.

An alternative means of computing the BSVD is suggested by the proof of Theorem 2. Aside from basic matrix manipulation, the technique indicated revolves around the singular value decompositions (2.2) and (2.3) and the orthogonalization (2.5). However, in some situations, not all these factorizations are necessary. This is implicit in Lawson and Hanson [7, p. 188ff.]. For example, if rank $(A)=n$, then U, V and X in (2.1) can be determined as follows:

1. $U_1^TAP=\begin{pmatrix}R\\0\end{pmatrix}$. (Householder triangularization with pivoting, $R\in R^{n\times n}$. See [7].)
2. $V^T(BPR^{-1})Z=D_B$. (SVD of BPR^{-1}.)
3. Set $X=PR^{-1}Z$ and $U=U_1\begin{pmatrix}Z & 0\\0 & I_{m_a-n}\end{pmatrix}$.

We now formulate a generalized singular value decomposition in connection with (1.2′). It is a simple combination of the standard SVD and the following definitions.

Definition 2. Let $P\in R^{n\times n}$ *be positive definite. A matrix* $Q\in R^{n\times n}$ is P orthogonal if $Q^TPQ=I_n$.

Definition 3. Let S and T be positive definite matrices of orders m and n, respectively, with $m\geqq n$. The S, T-singular values of $A\in R^{m\times n}$ are the elements of the set $\mu_{S,T}(A)$ defined by

$$\mu_{S,T}(A)=\left\{\mu\mid\mu\geqq 0,\ \mu \text{ a stationary value of } \frac{\|Ax\|_S}{\|x\|_T}\right\}$$

Theorem 3 (the S, T-singular value decomposition). *Let A, S and T be in $R^{m\times n}$, $R^{m\times m}$ and $R^{n\times n}$, respectively, with S and T positive definite $(m\geqq n)$. There exists an S orthogonal $U\in R^{m\times m}$ and a T orthogonal $V\in R^{n\times n}$ such that*

$$U^{-1}AV=D=\operatorname{diag}(\mu_1,\cdots,\mu_n)$$

Furthermore, $\mu_{S,T}(A)=\{\mu_1, \cdots, \mu_n\}$.

Proof. Let $S=LL^T$ and $T=KK^T$ be the Cholesky factorization of S and T [7]. Set $C=L^TAK^{-T}$ $(K^{-T}=(K^T)^{-1}=(K^{-1})^T)$ and let

$$Q^TCZ=D=\operatorname{diag}(\mu_1, \cdots, \mu_n)$$

represent its SVD. By defining $U=L^{-T}Q$ and $V=K^{-T}Z$, we see that U and V are S and T orthogonal, respectively, and that $U^{-1}AV=D$.

By using Lagrange multipliers we see that the stationary values of $\|Ax\|_S/\|x\|_T$ are precisely the zeros of $\det(A^TSA-\mu^2T)=0$. Since

$$\begin{aligned}\det(A^TSA-\mu^2T)&=\det(T)\det(C^TC-\mu^2I_n)\\&=\det(T)\det(D^TD-\mu^2I_n)\\&=\det(T)\prod_1^n(\mu_i^2-\mu^2),\end{aligned}$$

it is clear that $\mu_{S,T}(A)$ is correctly given by the theorem. Q.E.D.

There are some connections between the two generalized singular value decompositions presented thus far. For example, if $\operatorname{rank}(B)=n$ in Theorem 2, then $\mu(A,B)=\mu_{I_{m},B^TB}(A)$ and the matrix $X\operatorname{diag}(\beta_1^{-1}, \cdots, \beta_n^{-1})$ is B^TB orthogonal.

3. Problems where generalized singular value techniques are applicable.

(a) *Damped least squares.* There are occasions when one wants to minimize the quadratic form

$$\|Ax-b\|^2+\lambda^2\|Bx-d\|^2, \tag{3.1}$$

where $A\in R^{m_a\times n}$ $(m_a\geqq n)$, $B\in R^{m_b\times n}$, $b\in R^{m_a}$, $d\in R^{m_b}$ and $\lambda>0$. There exists a unique solution $\hat{x}=\hat{x}(\lambda)$ of minimal norm to this problem and it is often of interest to know how $\hat{x}$ varies with λ.

In their discussion of this problem, Lawson and Hanson diagonalized A and B and were thus, in effect, doing a generalized singular value analysis of (3.1). We shall reproduce this portion of their work using Theorem 2, but unlike their analysis, which presumed B to be square and invertible, our reduction of (3.1) will be valid for arbitrary B. Indeed, if we substitute (2.1) into (3.1), we obtain

$$\begin{aligned}\|Ax-b\|^2+\lambda^2\|Bx-d\|^2&=\|D_Ay-\tilde{b}\|^2+\lambda^2\|D_By-\tilde{d}\|^2\\&=\sum_{j=1}^n(\alpha_jy_j-\tilde{b}_j)^2+\lambda^2\sum_{j=1}^q(\beta_jy_j-\tilde{d}_j)^2,\end{aligned}$$

where $y=X^{-1}x$, $\tilde{b}=U^Tb$ and $\tilde{d}=V^Td$ have components y_i, $\tilde{b}_i$ and $\tilde{d}_i$. Among the y which minimize $\|D_Ay-\tilde{b}\|^2+\lambda^2\|D_B-\tilde{d}\|^2$, it is clear that $\hat{y}=(\hat{y}_1, \cdots, \hat{y}_n)^T$ has minimal norm where

$$\hat{y}_j=\frac{\alpha_j\tilde{b}_j+\lambda^2\beta_j\tilde{d}_j}{\alpha_j^2+\lambda^2\beta_j^2}$$

for $j = 1, \cdots, r = \text{rank}(B)$ and

$$\hat{y}_j = \begin{cases} \dfrac{\tilde{b}_j}{\alpha_j}, & \alpha_j \neq 0, \\ 0, & \alpha_j = 0, \end{cases}$$

for $j = r+1, \cdots, n$. Thus, $\hat{x} = X\hat{y}$ minimizes (3.1).

We still must show that $\hat{x}$ has minimal norm. Suppose x_i denotes the ith column of X. From Remark 3 after Theorem 2, we know that $x_1, \cdots, x_k$ are orthogonal to $NS(A) \cap NS(B)$. Hence, the same can be said of $\hat{x} = \sum_1^r \hat{y}_i x_i$ because $r = \text{rank}(B) \leqq \text{rank}\begin{pmatrix} A \\ B \end{pmatrix} = k$. From this one can deduce that $\hat{x}$ has minimal norm.

Having derived an equivalent, diagonal version of (3.1), one can vary λ so that the solution $\hat{y}$ (and hence, $\hat{x}$) assumes certain properties. We refer the reader to Lawson and Hanson [7].

(b) *Least squares with equality constraints.* Suppose $A \in R^{m_a \times n}$ $(m_a \geqq n)$, $B \in R^{m_b \times n}$, $b \in R^{m_a}$, $d \in R^{m_b}$ and the set $\{x \mid Bx = d\}$ is nonempty. We want to consider the problem

$$\min_{Bx=d} \|Ax = b\|. \tag{3.2}$$

By using the B-singular value decomposition (2.1), we transform this to

$$\min_{D_By=\tilde{d}} \|D_A y - \tilde{b}\|,$$

where $y = X^{-1}x$, $\tilde{b} = U^T b$ and $\tilde{d} = V^T d$. A solution $\hat{x}$ to (3.2) is thus given by $\hat{x} = X\hat{y}$, where $y = (\hat{y}_1, \cdots, \hat{y}_n)^T$ is defined by

$$\hat{y}_i = \frac{\tilde{d}_i}{\beta_i}, \qquad i = 1, \cdots, r,$$

$$\hat{y}_i = \begin{cases} 0, & \alpha_i = 0, \\ \dfrac{\tilde{b}_i}{\alpha_i}, & \alpha_i \neq 0, \end{cases} \qquad i = r+1, \cdots, n.$$

An argument similar to the one given in the previous application shows that $\hat{x}$ is the unique, minimal norm solution of (3.2).

Lawson and Hanson pay considerable attention to (3.2), and we refer the reader to their discussion of algorithms and theory. We remark in passing that their analysis in Chapter 22 can be extended by using Theorem 2. This enables one to dispense with the assumption $\text{rank}\begin{pmatrix} A \\ B \end{pmatrix} = n$ $\left(\text{rank}\begin{pmatrix} C \\ E \end{pmatrix} = n \text{ in their notation}\right)$ on page 152.

Leringe and Wedin [8] have done a perturbation analysis of (3.2) and have concluded that the problem is well-conditioned when the matrices AP and B are themselves well-conditioned. (Here the columns of $P \in R^{n \times (n-r)}$ form an orthonormal basis for $NS(B)$.) The B-singular value decomposition provides an explicit way of seeing this. Assume that the columns of X are denoted by x_i and that $\|x_i\| = 1$ for $i = 1, \cdots, n$. From

$$\hat{x} = \sum_{i=1}^{r} d_i\left(\frac{x_i}{\beta_i}\right) + \sum_{\substack{i=r+1 \\ \alpha_i \neq 0}}^{n} b_i\left(\frac{x_i}{\alpha_i}\right)$$

we see that the sensitivity of the solution depends upon the reciprocals of $\beta_1, \cdots, \beta_r$ and $\alpha_{r+1}, \cdots, \alpha_n$. If β is the smallest singular value of B and α is the smallest singular value of AP, then one can use Remark 4 after Theorem 2 to show that $\beta_i \geqq \beta$, $i = 1, \cdots, r$, and $\alpha_i \geqq \alpha$, $i = r+1, \cdots, n$. Hence, if B and AP are well-conditioned, the reciprocals of the relevant β_i and α_i are modest. In this case, $\hat{x}$ will not be too sensitive to changes in the vectors $\tilde{b} = U^T b$ and $\tilde{d} = V^T d$.

(c) *Certain generalized eigenvalue problems.* Let $\{f_k\}_1^n$ be a family of "basis functions" on $[0, 1]$ and let $p(x) \geqq \delta > 0$ be defined on $[0, 1]$. Consider the matrix $C = (c_{ij}) \in R^{n \times n}$ where

$$c_{ij} = \int_0^1 p(x) f_i(x) f_j(x)\, dx. \tag{3.3}$$

If we evaluate these integrals using a quadrature rule with abscissas $\{x_k\}_1^m$ and positive weights $\{\omega_k\}_1^m$, then

$$C \approx \tilde{C} = (\tilde{c}_{ij}) = A^T A,$$

where $c_{ij} \cong \tilde{c}_{ij} = \sum_{k=1}^m \omega_k p(x_k) f_i(x_k) f_j(x_k)$, $A = (a_{ij}) \in R^{m \times n}$ and $a_{ij} = \sqrt{\omega_i p(x_i)} \cdot f_j(x_i)$.

This observation has a bearing upon certain generalized eigenvalue problems $Gx = \lambda Hx$ when the entries in G and H are integrals of the form (3.3). Such problems arise when Rayleigh–Ritz procedures are applied to the differential equation $-D[p(x)Du(x)] = \lambda u(x)$ where $0 \leqq x \leqq 1$ and $u(0) = u(1) = 0$. (See [2].) In this case, the original $Gx = \lambda Hx$ problem is approximated through quadrature by a B-singular value problem $A^T A x = \lambda B^T B x$. Since formation of the products $A^T A$ and $B^T B$ may induce errors, and since the matrices A and B are so simple to construct, there may be reason to apply generalized singular value techniques to A and B rather than generalized eigenvalue techniques to $A^T A$ and $B^T B$. The development of sparse generalized singular value algorithms would make research in this direction particularly fruitful. Some recent work by Argyris and Bronlund [1] fortifies this point of view. Their "natural factor" formulation of certain stiffness matrices enables one to avoid the inaccuracies which attend products like $A^T A$.

(d) *Weighted least squares.* Let $A \in R^{m \times n}$ $(m \geqq n)$, $b \in R^m$, $S \in R^{m \times m}$ (positive definite), and $T \in R^{n \times n}$ (positive definite). Consider the problem of minimizing

$$\|Ax - b\|_S$$

such that

$$\|x\|_T = \min.$$

The solution $\hat{x}$ to this problem is unique. If $U^{-1} A V = D_A = \operatorname{diag}(\mu_1, \cdots, \mu_n)$ is the S, T-singular value decomposition of A $(\mu_1 \geqq \cdots \geqq \mu_r > \mu_{r+1} = \cdots = \mu_n = 0)$, then it can be shown that $\hat{x} = VD_A^+ U^{-1} b$, where $D_A^+ = \text{diag}(1/\mu_1, \cdots, 1/\mu_r, 0, \cdots, 0) \in R^{n \times m}$.

The case when S and T are diagonal is of special interest, for then a "fast Givens" version of the Golub–Reinsch SVD routine becomes particularly appropriate. We refer the reader to Van Loan [12] for a discussion of this algorithm and to Gentleman [3] for a description of these modified Givens techniques relative to weighted least squares problems.

We briefly summarize what the algorithm developed in [12] does. Given A and the positive diagonals S and T, it finds invertible $U_1(m \times m)$ and $V_1(n \times n)$ such that (i) $U_1^T A V_1 = D$ is nonnegative diagonal, (ii) $U_1^T S^{-1} U_1 = S_1$ is positive diagonal and (iii) $V_1^T T V_1 = T_1$ is also positive diagonal. This implies that $U = U_1^{-T} S^{1/2}$ is S orthogonal and that $V = V_1 T^{-1/2}$ is T orthogonal. Thus $D_a \equiv S_1^{-1/2} D T_1^{-1/2} = S_1^{-1/2} U_1^t A V_1 T_1^{-1/2} = U^{-1} AV$ is the S, T-singular value decomposition of A and furthermore, $\hat{x} = V_1 T^{-1/2} T_1^{1/2} D^+ S_1^{1/2} S^{-1/2} U_1^T b$. Notice that the only inverses in the expression for $\hat{x}$ involve diagonal matrices.

Acknowledgments. Some of this work was part of the author's Ph.D. thesis written at the University of Michigan under Cleve B. Moler, whom the author wishes to thank. In addition, the author is obliged to the referee for his many valuable comments.

REFERENCES

[1] J. H. ARGYRIS AND O. E. BRONLUND, *The natural factor formulation of the stiffness for the matrix displacement method*, Computer Methods in Applied Mechanics and Engineering, 5 (1975), pp. 97–119.

[2] G. FIX AND G. STRANG, *An Analysis of the Finite Element Method*, Prentice-Hall, Englewood Cliffs, N.J., 1973.

[3] W. M. GENTLEMAN, *Least squares computations by Givens transformations without square roots*, J. Inst. Math. Appl., 12 (1973), pp. 329–336.

[4] G. H. GOLUB AND P. BUSINGER, *An ALGOL procedure for computing the singular value decomposition*, Stanford University Rep. CS73, Stanford Univ., Stanford, Calif., 1967.

[5] G. H. GOLUB AND C. REINSCH, *Singular value decomposition and least squares solutions*, Numer. Math., 14 (1970), pp. 403–420.

[6] W. KAHAN AND G. H. GOLUB, *Calculating the singular values and pseudoinverse of a matrix*, this Journal, 2 (1965), pp. 205–224.

[7] C. LAWSON AND R. HANSON, *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, N.J., 1974.

[8] O. LERINGE AND P. WEDIN, *A comparison between different methods to compute a vector x which minimizes $|Ax - b\|_2$ when $Gx = h$*, Lund University Computer Science Rep., Lund Univ., Lund, Sweden, 1970.

[9] C. B. MOLER AND G. W. STEWART, *An algorithm for generalized matrix eigenvalue problems*, this Journal, 10 (1973), pp. 241–256.

[10] G. W. STEWART, *On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$*, this Journal, 9 (1972), pp. 669–686.

[11] C. F. VAN LOAN, *A general matrix eigenvalue algorithm*, this Journal, 12 (1975), pp. 819–834.

[12] ———, *Generalized singular values with algorithms and applications*, Ph.D. thesis, University of Michigan, Ann Arbor, Michigan, 1973.