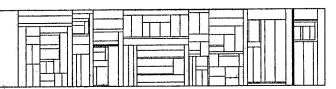
METHODS FOR UPDATING THE SINGULAR VALUE DECOMPOSITION

bу

Linda Kaufman

DAIMI PB-26 March 1974

Institute of Mathematics University of Aarhus
DEPARTMENT, OF COMPUTER SCIENCE
Ny Munkegade - 8000 Aarhus C - Denmark
Phone 06-128355



METHODS FOR UPDATING THE SINGULAR VALUE DECOMPOSITION Linda Kaufman

Abstract

The linear least squares problem of minimizing $\|Ax - b\|\|_2$ where A is an m×n matrix, m \geq n, may be solved using the singular value decomposition in approximately $2mn^3 + 4n^3$ multiplications. In this paper the problem of solving $\|A'x - b\|\|_2$ is considered where A' results from deleting or adding a column to A. This might occur when a change is made in the model of a process. Instead of computing the singular value decomposition of A' from scratch, the singular value decomposition of A is updated. Since the updating require about $6n^3$ multiplications the algorithms are useful when m>>n. The problem of recalculating some or all of the singular values of a matrix A', which is obtained by deleting or adding a row or a column from a matrix A, whose singular value decomposition is known, is also studied.

Section 1. Introduction

The singular value decomposition (see Rao and Mitra [8]) of an $m \times n$ matrix A of rank r is given by

$$(1.1) A = U \Sigma V^{T}$$

where $U(m \times m)$ and $V(n \times n)$ are orthogonal matrices and Σ is an $m \times n$ diagonal matrix with elements $(\sigma_1, \sigma_2, \dots, \sigma_r, 0, 0, 0)$ where $\sigma_{i-1} \geq \sigma_i$. The σ 's are called the singular values of A and are the positive square roots of the eigenvalues of A^TA . The decomposition (1.1) has many applications (see [5] and [8]), perhaps the best known of which is the solution of the linear least squares problem in which given the matrix A and a vector Σ 0 one finds the vector Σ 0 of minimal length which minimizes

$$||A \times -b||_2$$
.

Since $\|Ax - b\|_2 = \|U\Sigma V^T x - b\|_2 \| = \|\Sigma V^T x - U^T b\|_2$, once can easily show that

$$x_{O} = V\Sigma^{+} U^{T} b$$

where the pseudoinverse Σ^+ is given by

$$\Sigma^{+} = \begin{pmatrix} \sigma_{1}^{-1} & 0 & \\ & \sigma_{1}^{-1} & \\ & & \sigma_{r}^{-1} & \\ & & & 0 & \end{pmatrix}$$

Golub and Reinsch [4] have given a numerically stable algorithm which computes $\underset{\sim}{x_0}$ in about $(2 \text{ m n}^3 + 4 \text{ n}^3)$ multiplications and additions. Businger [2] has given algorithms which update $\overset{x}{\sim}_{0}$ when a new row is added or deleted from the system without recomputing the new singular value decomposition from scratch. His algorithm requires about $6n^3$ multiplications and additions and hence is useful when m >> n, the most common case in statistical applications. In Sections 2 and 3 of this paper we will present algorithms for updating \mathbf{x} when a column is deleted from or added to A. This might occur when there is a decision to change the model for a process. Like Businger's schemes, the procedures are practical when m >> n. The algorithm for adding a column is quite similar to Businger's for adding a row, but the algorithm for deleting a column is more similar to the approaches given in Gill, Golub, Murray, and Saunders [4] which unlike Businger's scheme, avoids complex arithmetic for real matrices.

Section 2. Adding a column

Let A' = [A; a] and assume the singular value decomposition of A is given by

$$A = U \Sigma V^{T}.$$

If P and Q are orthogonal matrices, then solving

is equivalent to finding x_1 which minimizes

$$\text{HP}(\Sigma: U^{T}\underline{a})QQ^{T}V^{T}\underline{x} - PU^{T}\underline{b}||_{2}^{2}.$$

If P and Q are determined such that $P^{T}(\Sigma:U^{T}\underline{a})Q$ is a diagonal matrix Σ' , then

$$x_1 = V Q \Sigma'^+ P^T c$$
 where $c = U^T b$.

The algorithm for finding Σ' consist of 2 phases

- (1) The reduction of $L = (\Sigma: U^{T}a)$ to a bidiagonal matrix J using plane rotations as explained below.
- (2) The reduction of J to Σ' using the second part of the Golub-Reinsch algorithm.

The matrix L has the form

The first phase begins with the formation of the Householder transformation $P_0 = I - uu^T$ which zeros $l_{i,n+1}$ for i > n+1 when applied to L. Thus the matrix $L' = P_0L$ has the form

The transformation P_o is also applied to c. Reducing L' to J requires n major steps, of which the k^{th} zeroes $l_{k,n+1}^{'}$. Each step consists of applying to L' a number of Givens transformations R_i , of the form

where $c^2 + s^2 = 1$. We show the k^{th} step which is typical. At the beginning of the k^{th} step, L' has the form

(2.1)
$$\begin{array}{cccc}
\times \times \times & 0 \\
\times \times \times & 0 \\
\times & \times & \times \\
\times & \times \\
\times & \times
\end{array}$$

A Givens rotation R_k is chosen so that the (k,n+1)th element of R_kL' is zero. The matrix R_kL' has the form

The unwanted (k+1,k)th element is chased up the diagonal using column and row rotations in the following steps:

- ⊗ eliminated element,
- + introduced nonzero element.

The main problem with this algorithm is the necessity of having the U matrix handy. Unfortunately the Golub-Reinsch algorithm does not construct U or store sufficient information for its construction. To modify their algorithm to save U would be quite costly. However, the situation is not that bleak. In the first part of the Golub-Reinsch algorithm the orthogonal matrices Q and Z are determined such that

$$A = Q J Z$$

where J is bidiagonal. The matrix Q is a product of Householder transformations, and although it is not explicitly constructed, it (and its transpose) can be easily reconstructed. The information for its reconstruction can be stored in the "Zeroed" portion of the A matrix. If the decomposition (2.3) replaces the singular value decomposition of A, then our algorithm is exactly the same with $V^{\rm T}$ replaced by Z and L given by

At the beginning of the $\,k^{\mbox{th}}\,$ step of the algorithm L' would have the form

instead of that given in (2.1). The same sequence of Givens rotations would reduce this matrix to bidiagonal form.

Section 3. Deleting a column

Assume that the singular decomposition is given by

$$A = U \Sigma V^{T}$$

and that it is desired to find the minimal solution $\underset{\sim}{\mathbf{x}_1}$ of

where A' is formed by dropping one column, without loss of generality the last column, from A.

If orthogonal matrices Q and P can be found such that

$$(3.2) VQ = \left(\frac{V' \mid Q}{Q^T \mid 1} \right)$$

and PSQ = ${\rm diag}(\bar{\sigma}_i) = \bar{\Sigma}$ and the matrix U' is set to UP^T and Σ' to $\bar{\Sigma}$ E_{n-1} , where E_{n-1} is the first n-1 columns of the identity matrix, then because

$$\overline{\Sigma} = PU^TAVQ = PU^T[A': \underline{a}] \left(\frac{V'|0}{0|1}\right) = (PU^TA'V'|PU^T\underline{a}),$$

it should be obvious that

$$A' = U' \Sigma' V'^{T}$$

is the singular value decomposition of A'. Thus the solution of (3.1) is given by

$$x_1 = V' \Sigma'^+ U'^T b$$
.

The problem reduces to determining P and Q which satisfy (3.2). As in Section 1, the algorithm has 2 stages. In the first stage orthogonal matrices P_a and Q_a are found such that

and $(P_a \Sigma Q_a) E_{n-1} = J$ an mx(n-1) bidiagonal matrix. Equation (3.3) implies that there exist a matrix V_i and a vector \underline{q} such that

$$v Q_a = \left(\frac{v_1}{o} \middle| \frac{q^T}{a}\right),$$

but since VQ_a is orthogonal, $g^T = Q^T$.

In the second stage the matrix J is reduced to diagonal form by the second part of the Golub-Reinsch algorithm using row rotations $P_{\rm b}$ and column rotations $Q_{\rm b}$. The matrices Q and P are simply

$$Q = Q_a \left(\frac{Q_b}{0} \right) \quad \text{and} \quad P = P_b P_a$$

Determining P_a, Q_a and J requires n-1 major steps. If P_k and Q_k represent the product of row and column transformations up until the k^{th} step, then the k^{th} step is designed to zero the k^{th} element of $\chi^{(k)T} \equiv \chi_n^T Q_K$ and to keep the bidiagonality of

$$J_k = P_k \Sigma Q_k$$

At the beginning of the k^{th} step, $\mathbf{v}^{(k)T}$ looks like

$$(0 \cdots 0 , \times \times \times \times)$$

and J_k looks like

A column Givens rotation R_k is chosen so that the k^{th} element of $\mathbf{v}^{(k)}\mathbf{T}R_k$ is zero. When R_k is applied to \mathbf{J}_k , the

result is a matrix whose form is

To get rid of the unwanted (k+1,k)th element a row transformation R_k is applied. This is followed by another row transformation to zero the (k-1,k+1)st element. The result is a matrix which looks like

The unwanted (k,k-1) element is chased up the diagonal as in (2.2). Since none of the column transformations in the chasing process affect the k+1st plane, the zeroes in $v_k^{(k)}$ R_k are not disturbed.

Section 4. Updating the singular values

In this section we will consider methods for determining some or all of the singular values of a matrix \overline{A} which differs from another matrix A, whose singular value decomposition is given by $A = U \Sigma V^T$, by the deletion or addition of a row or column.

Finding a few of the singular values via the Sturm Sequence Algorithm

1) Adding a row $\bar{\bar{A}} = \begin{bmatrix} \frac{a}{\sim} \\ \bar{\lambda} \end{bmatrix} .$

Since

$$\left(\frac{1 \mid 0}{0 \mid \mathbf{U}^{\mathrm{T}}} \right) \left[\frac{\mathbf{a}^{\mathrm{T}}}{\mathbf{A}} \right] \left[\mathbf{a} \mid \mathbf{A}^{\mathrm{T}} \right] = \left[\frac{1 \mid 0}{0 \mid \mathbf{U}} \right] = \left(\mathbf{a}^{\mathrm{T}} \mathbf{a} \quad \mathbf{a}^{\mathrm{T}} \mathbf{A}^{\mathrm{T}} \mathbf{U} \right) \equiv \mathbf{D},$$

the singular values of \overline{A} are the positive square roots of the nonzero eigenvalues of D. Since the matrix D has the form

$$\begin{pmatrix} \frac{d_1 & b_1^T}{b_1} \\ \frac{d_2 \cdot d_{m+1}}{b_1} \end{pmatrix}, \quad b_1^T = (b_2, \dots, b_n)$$

one can easily determine its eigenvalues by the Sturm Sequence algorithm (see Wilkinson and Reinsch [9]) which requires the signs of the principal minors of $D-\alpha I$ for various values of α .

If one sets

$$c_{0} = 1$$
 , $\bar{d}_{i} = d_{i} - \alpha$, $e_{0} = 1$, $b_{1} = 0$ and $e_{i} = -\bar{d}_{i}e_{i-1} + b_{i}^{2}c_{i-1}$ $c_{i} = c_{i-1}\bar{d}_{i}$

then it is easy to show by expanding $\,D\text{-}\alpha I\,\,$ by its i $^{\rm th}$ row that $e_{\,i}\,\,$ contains the i $^{\rm th}$ principal minor.

2) Deleting a row

$$A = \left(\frac{\overline{A}}{a^{T}}\right)$$

If an orthogonal matrix P is constructed such that $P(U^T \in_{m}) = e_{m}$, then because of the orthogonality of the two matrices PU^T has the form

$$\left(\frac{\overline{U} + Q}{0^{\mathrm{T}} + 1}\right)$$

Thus

which means that the singular values of $\bar{\mathbf{A}}$ are the positive square roots of the eigenvalues of

$$(4.1) E_{m-1}^{T} P \Sigma \Sigma^{T} P^{T} E_{m-1}$$

These values can be calculated very easily with a Sturm Sequence approach if P is chosen correctly. The P matrix can be written as

$$P = R_{m-1}, \dots, R_1$$

where R_k is designed to zero the k^{th} element of $R_{k-1}\cdots R_1(U^Te_m)$ and has the form

and $c_k^2 + s_k^2 = 1$. Given a diagonal matrix $D = diag(d_i)$ the k^{th} principal minor of PDP^T is the k^{th} principal minor of

$$R_k \dots R_1 DR^T \dots R_k^T$$

and the kth principal minor of

$$R_{k-1} \dots R_1 DR_1^T \dots R_{k-1}^T = \prod_{i=1}^k d_i$$

The matrix

$$R_{k-1} \cdots R_1 DR_1^T R_{k-1}^T$$

has the form

$$k^{\text{th}} \text{ row} \left(\begin{array}{c|c} x & a \\ \hline a & y & d_{k+1} \\ \hline & & \cdot d_n \end{array} \right)$$

If R_k is applied to its left and right, its k^{th} principal submatrix would be

$$\det\begin{pmatrix} x & c_k \\ c_k a & c_k^2 y + s_k^2 d_{k+1} \\ & & + c_k^2 & \prod_{i=1}^{k} d_i \end{pmatrix}$$

Since det(X) is the k-1st principal minor, we can write the following recursive for the k^{th} principal minor e_k :

$$e_{o} = 0,$$
 $f_{o} = 1,$
 $f_{k} = d_{k}f_{k-1}$
 $e_{k} = s_{k}^{2} d_{k+1}e_{k-1} + c_{k}^{2}f_{k}$

In our case we need the princpal minors of

$$E_{m-1}^{T} P \Sigma \Sigma^{T} P^{T} E_{m-1} - \alpha E_{m-1}^{T} E_{m-1} =$$

$$E_{m-1}^{T} P (\Sigma \Sigma^{T} - \alpha I) P^{T} E_{m-1}$$

so that $d_i = \sigma_i^2 - \alpha$ for $i \le r$ and $d_i = -\alpha$ for i > r.

3) Adding a column

$$\bar{A} = (a! A).$$

Since

$$\left(\frac{1 \mid 0}{0 \mid V^{T}}\right) \left(\frac{\overset{a}{\overset{T}{\overset{}}{\overset{}}{\overset{}}}}{\overset{A^{T}}{\overset{}}}\right) \quad (\overset{\sim}{\mathsf{a}} \mid \mathsf{A}) \quad \left(\frac{1 \mid 0}{0 \mid \mathsf{V}}\right) = \quad \left(\overset{\overset{a}{\overset{}}{\overset{}}{\overset{}}{\overset{}}}{\overset{a}{\overset{}}{\overset{}}} \frac{\overset{a}{\overset{}}{\overset{}}}{\overset{a}{\overset{}}} \frac{\overset{a}{\overset{}}{\overset{}}}{\overset{a}{\overset{}}} \frac{\mathsf{A}^{\mathsf{V}}}{\mathsf{A}^{\mathsf{T}}}\right) \quad = \quad \mathsf{D}$$

the singular values of \overline{A} are the square roots of the eigenvalues of D and one may proceed as if one were adding a row.

4) Deleting a column

$$A = (\bar{A}: \underline{a})$$

If an orthogonal matrix P is constructed such that $P \ V^T \ \underset{\sim}{e_n} \ = \ \underset{\sim}{e_n} \ , \ \text{then} \qquad \qquad \underset{\sim}{e_n^T} \ P U^T \ = \ \underset{\sim}{e_n^T}$

and because of the orthogonality of the 2 matrices, PV^{T} has the form

$$\begin{pmatrix} \overline{V} & 0 \\ 0 & 1 \end{pmatrix}$$
 .

Thus

$$P \Sigma^{T} \Sigma P = P V^{T} \begin{pmatrix} \overline{A}^{T} \\ \underline{a}^{T} \end{pmatrix} (\overline{A}: \underline{a}) V P^{T}$$

$$= \left(\frac{\overline{V} \mid 0}{0 \mid 1} \right) \left(\frac{\overline{A}^{T}}{a^{T}} \right) (\overline{A}: \underline{a}) \left(\frac{\overline{V}^{T} \mid 0}{0 \mid 1} \right)$$

$$= \left(\frac{\overline{V} \overline{A}^{T} \overline{A} \overline{V} \mid \overline{V} \overline{A}^{T} \underline{a}}{a^{T} \overline{A} \overline{V}^{T} \mid \overline{a}^{T} \underline{a}} \right)$$

and hence one may proceed as if one were deleting a row.

Finding all the eigenvalues by a Generalized Eigenvalue Problem

1) Adding a row. Let
$$A = U \Sigma V^T$$
 and $\overline{A} = \begin{bmatrix} \overset{\cdot}{a}^T \\ A \end{bmatrix}$. Since $A^T A = V \Sigma^T \Sigma V$,
$$\overset{\cdot}{V}^T (\overset{\cdot}{a} \colon A^T) (\frac{A}{\overset{\cdot}{a}^T}) V = \Sigma^T \Sigma + (V^T \overset{\cdot}{a}) (V^T \overset{\cdot}{a})^T$$
$$= D + \overset{\cdot}{b} \overset{\cdot}{b}^T ,$$

which means that the singular values of \bar{A} are the square roots of the eigenvalues of $D + b b^T$. Bartels, Golub, and Saunders [1] have shown that solving the eigenvalue problem

$$(D + b b^{T}) x = \lambda x$$

is equivalent to solving the generalized eigenvalue problem

$$(4.2) A \overset{\cdot}{\times} = \lambda B \overset{\cdot}{\times}$$

where A and B are both symmetric and tridiagonal and B is positive definite and

$$A = K D K^{T} + K \underset{\sim}{b} \underset{\sim}{b}^{T} K^{T}$$

$$B = K K^{T}$$

where K is the bidiagonal matrix such that

$$K \underset{b}{\triangleright} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Crawford [3] and Peters and Wilkinson [7] have recently given methods for solving (4.2).

2) Deleting a row.

Let
$$A = (\bar{A}_T)$$
.

Since

$$\Sigma^{T}\Sigma = V^{T}(\overline{A}^{T}: \underline{a})(\underline{\overline{A}}^{T})V = V^{T}\overline{A}AV + (V^{T}\underline{a})(V^{T}\underline{a})^{T}$$

the singular values of $ar{\mathtt{A}}$ are the square roots of the eigenvalues of

$$D - b b^T$$

where $D = \Sigma^{T}\Sigma$ and $b = V^{T}a$. As in the previous case these

eigenvalues can be found by solving a generalized eigenvalue problem like (4.2). In this case

(4.3)
$$A = KDK^{T} - (K b) (K b)^{T}.$$

3) Adding a column.

Let
$$\bar{A} = (a: A)$$
.

Since

$$\mathbf{U}^{\mathrm{T}}(\mathbf{a}: \mathbf{A}) \begin{pmatrix} \mathbf{a}^{\mathrm{T}} \\ \mathbf{\Delta}^{\mathrm{T}} \end{pmatrix} \mathbf{U} = (\mathbf{U}^{\mathrm{T}}\mathbf{a}) (\mathbf{U}^{\mathrm{T}}\mathbf{a})^{\mathrm{T}} + \Sigma \Sigma^{\mathrm{T}},$$

the singular values of \bar{A} are the square roots of the nonzero eigenvalues of problem (4.2), with $D = \Sigma \Sigma^{T}$ and $b = U^{T}a$.

4) Deleting a column.

Let
$$A = (a: \overline{A})$$
.

Since

$$\Sigma \Sigma^{\mathrm{T}} = \mathbf{U}^{\mathrm{T}} (\mathbf{a} : \mathbf{A}) (\mathbf{a}^{\mathrm{T}}) \mathbf{U} = \mathbf{U}^{\mathrm{T}} \mathbf{A} \mathbf{A}^{\mathrm{T}} \mathbf{U} + \mathbf{U}^{\mathrm{T}} \mathbf{a} \mathbf{a}^{\mathrm{T}} \mathbf{U}$$

the singular values of \bar{A} are the square roots of the nonzero eigenvalues of (4.3) with $D = \Sigma \Sigma^{T}$ and $b = U^{T}a$.

We note that among the algorithms just presented there exists an algorithm for finding the singular values of a matrix, which has resulted from the deletion from or addition to a row or a column of another matrix A, without requiring the U matrix of the singular decomposition of A. Thus the Golub-Reinsch algorithm, which does not compute U, can be used without modification.

REFERENCES

- [1] R.H. Bartels, G.H. Golub, and M.A. Saunders, "Numerical Techniques in Mathematical Programming", Nonlinear Programming, Academic Press, New York, (1970).
- [2] P.A. Businger, "Updating a Singular Value Decomposition", BIT 10 (1970) 376-397.
- [3] C.R. Crawford, "Reduction of a Band-Symmetric Generalized Eigenvalue Problem", Communications of the ACM 16 (1973), 41-44.
- [4] P.L.Gill, G.H. Golub, W. Murray, and M.A. Saunders,
 "Methods" for Modifying Matrix Factorizations",
 Stanford Computer Science Report 322, (November 1972).
- [5] G.H. Golub, "Matrix Decompositions and Statistical Calculations: In R.C. Milton and J.A. Helder (editors), Statistical Computation, Academic Press, New York (1969).
- [6] G.H. Golub and C. Reinsch, "Singular Value Decomposition and Least Squares Solutions", Numer.Math. 14 (1970), 403-420.
- [7] G. Peters and J.H. Wilkinson, " $Ax = \lambda Bx$ and the Generalized Eigenproblems", SIAM J. Numer.Anal. 7 (1970), 479-492.
- [8] C.R. Rao and S.K. Mitra, Generalized Inverse of Matrices and Its Applications , Wiley, New York (1971).
- [9] J.H. Wilkinson and C. Reinsch, <u>Handbook for Automatic</u>

 <u>Computation</u>, Springer Verlag, New York (1971).