METHODS FOR UPDATING
THE SINGULAR VALUE DECOMPOSITION

by

Linda Kaufman

DAIMI PB-26
March 1974
METHODS FOR UPDATING THE SINGULAR VALUE DECOMPOSITION

Linda Kaufman

Abstract

The linear least squares problem of minimizing $\|Ax - b\|_2$ where $A$ is an $m \times n$ 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 $m \gg n$. 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

\[
A = U \Sigma V^T
\]

(1.1)

where \( U(m \times m) \) and \( V(n \times n) \) are orthogonal matrices and \( \Sigma \) is an \( m \times n \) diagonal matrix with elements \( \sigma_1, \sigma_2, \ldots, \sigma_r, 0, 0, 0 \) where \( \sigma_{i-1} \geq \sigma_i \). The \( \sigma \)'s are called the singular values of \( A \) and are the positive square roots of the eigenvalues of \( A^T A \). 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 \( b \) one finds the vector \( x_0 \) of minimal length which minimizes

\[
\| A x - b \|_2.
\]

Since \( \| A x - b \|_2 = \| U \Sigma V^T x - b \|_2 = \| \Sigma V^T (x - U^T b) \|_2 \), once can easily show that

\[
x_0 = V \Sigma^+ U^T b
\]

where the pseudoinverse \( \Sigma^+ \) is given by

\[
\Sigma^+ = \begin{pmatrix}
\sigma_1^{-1} & 0 \\
0 & \ddots & 0 \\
0 & \cdots & \sigma_r^{-1}
\end{pmatrix}
\]
Golub and Reinsch [4] have given a numerically stable algorithm which computes $\xi_0$ in about $(2m n^3 + 4n^3)$ multiplications and additions. Businger [2] has given algorithms which update $\xi_0$ 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 $m \gg n$, the most common case in statistical applications. In Sections 2 and 3 of this paper we will present algorithms for updating $\xi$ 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 $m \gg n$. 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 $A' = [A; \bar{a}]$ 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

$$||A' \xi - b||_2$$

is equivalent to finding $\xi_1$ which minimizes

$$\min_{\Sigma: U_{\bar{a}}^T Q Q^T V^T \xi - P U^T \bar{b}} \|_2.$$
If $P$ and $Q$ are determined such that $P^T(Σ: U^T_{∞})Q$ is a diagonal matrix $Σ'$, then

$$
Σ_1 = VQΣ_1^+P^T_C \quad \text{where} \quad Q = U^T_C.
$$

The algorithm for finding $Σ'$ consist of 2 phases

1. The reduction of $L = (Σ: U^T_{∞})$ 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

```
  x  x  x  x
  x  x  x  x
  x  x  x  x
  x  x  x  x
```

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

```
  x  x  x  x
  x  x  x  x
  x  x  x  x
  x  x  x  x
```

The transformation $P_0$ is also applied to $Q$. 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 $R^*_l$, of the form

$$
\begin{bmatrix}
1 & c & s \\
-s & c & 1
\end{bmatrix}
$$

- $i$th row
where \( c^2 + s^2 = 1 \). We show the \( k \text{th} \) step which is typical.

At the beginning of the \( k \text{th} \) step, \( L' \) has the form

\[
\begin{array}{c}
\times & \times & 0 \\
\times & \times & 0 \\
\times & \times & + k \text{th row} \\
\times & \times & \\
\times & \\
\times & \\
\end{array}
\]

(2.1)

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

\[
\begin{array}{c}
\times & \times & 0 \\
\times & \times & 0 \\
\times & \times & 0 \\
\times & \times & \times \\
\times & \times & \\
\times & \\
\end{array}
\]

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

\[
\begin{array}{ccc}
\times & \times & 0 & \times & \times & 0 & \times & \times + 0 \\
\times & \times + 0 & \times & \times \times 0 & \times & \times 0 \\
\times & \times 0 & \times & \times \times \times 0 & \times & \times 0 \\
\times & \times \times \times \times \times \times \times 0 & \times & \times 0 \\
\times & \times 0 & \times & \times 0 & \times & \times 0 \\
\times & \times 0 & \times & \times 0 & \times & \times 0 \\
\times & \times 0 & \times & \times 0 & \times & \times 0 \\
\end{array}
\]

(1) \hspace{2cm} (2) \hspace{2cm} (3)

(2.2)

\[
\begin{array}{ccc}
\times & \times \times \times 0 & \times & \times \times 0 \\
+ \times & \times \times 0 & \times & \times \times 0 \\
\times & \times 0 & \times & \times \times 0 \\
\times & \times 0 & \times & \times \times 0 \\
\times & \times 0 & \times & \times \times 0 \\
\times & \times 0 & \times & \times \times 0 \\
\times & \times 0 & \times & \times \times 0 \\
\end{array}
\]

(4) \hspace{2cm} (5)

\( \times \) 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

\begin{equation}
A = QJZ
\end{equation}

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 $V^T$ replaced by $Z$ and $L$ given by

\begin{equation}
(J: Q A)
\end{equation}

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

\begin{align*}
\begin{array}{c|c}
\times & 0 \\
\times & 0 \\
\times & \times \\
\times & \times \\
\times & \times \\
\times & \\
\end{array}
\end{align*}

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 \( \hat{x}_1 \) of

\[ \| A' x - b \|_2 \]

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

\[ VQ = \begin{pmatrix} V' & 0 \\ \tilde{Q}^T & 1 \end{pmatrix} \]

and \( P \Sigma Q = \text{diag}(\tilde{\sigma}_1) = \tilde{\Sigma} \) and the matrix \( U' \) is set to \( U' = \tilde{U} \) and \( \Sigma' \) to \( \tilde{\Sigma} E_{n-1} \), where \( E_{n-1} \) is the first \( n-1 \) columns of the identity matrix, then because

\[ \tilde{\Sigma} = P U^T A V Q = P U^T [N; 0] \begin{pmatrix} V' & 0 \\ 0 & 1 \end{pmatrix} = (P U^T A V' \mid P U^T 0), \]

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

\[ \hat{x}_1 = V' \Sigma' + U'^T b. \]

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

\[ V_n^T Q_a = e_n^T \]
and \((P_a \Sigma Q_a)E_{n-1} = J\) an \(m \times (n-1)\) bidiagonal matrix. Equation (3.3) implies that there exist a matrix \(V\) and a vector \(q\) such that
\[
VQ_a = \begin{pmatrix} V_1 & q^T \\ 0 & 1 \end{pmatrix},
\]
but since \(VQ_a\) is orthogonal, \(q^T = q^T\).

In the second stage the matrix \(J\) is reduced to diagonal form by the second part of the Golub-Reinsch algorithm using row rotations \(P_b\) and column rotations \(Q_b\). The matrices \(Q\) and \(P\) are simply
\[
Q = Q_a \begin{pmatrix} Q_b & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad P = P_bP_a
\]

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 \(\hat{y}^{(k)T} = y^{T}_{n}Q_k\) and to keep the bidiagonality of \(J_k = P_k \Sigma Q_k\).

At the beginning of the \(k\)th step, \(\hat{y}^{(k)T}\) looks like
\[
(0 \ldots 0, x \times \times \times)
\]
and \(J_k\) looks like
\[
\begin{pmatrix}
\times & \times & \times & \times & \times & \times & k\text{th column} \\
\times & \times & \times & \times & \times & \times & \\
\times & \times & \times & \times & \times & \times & \\
\times & \times & \times & \times & \times & \times & \\
\times & \times & \times & \times & \times & \times & \\
\end{pmatrix}
\]

A column Givens rotation \(R_k\) is chosen so that the \(k\)th element of \(y^{(k)T}R_k\) is zero. When \(R_k\) is applied to \(J_k\), the
result is a matrix whose form is

\[
\begin{pmatrix}
\times & \times & \times & \cdots & \times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\end{pmatrix}
\]

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

\[
\begin{pmatrix}
\times & \times & \times & \cdots & \times & \times & \times & \cdots & \times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\times & \times & \times & \cdots & \times \\
\end{pmatrix}
\]

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 \(k+1\)st plane, the zeroes in \(y^{(k)}T\) \(R_k\) 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 \(\bar{A}\) 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
\[ \bar{A} = \begin{bmatrix} \bar{a}^T \\ \bar{a} \end{bmatrix} \]

Since
\[ \begin{pmatrix} 1 & 0 \\ 0 & U^T \end{pmatrix} \begin{bmatrix} \bar{a}^T \\ \bar{a} \end{bmatrix} \begin{bmatrix} A^T \\ 0 \end{bmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & U \end{pmatrix} = \begin{pmatrix} \bar{a}^T & \bar{a}^T A^T U \\ U^T \bar{a} \Sigma^T \end{pmatrix} = D, \]

the singular values of \( \bar{A} \) are the positive square roots of the nonzero eigenvalues of \( D \). Since the matrix \( D \) has the form
\[ \begin{pmatrix} d_1 & b_1^T \\ b_2 & 0 \\ \vdots & \ddots & \ddots \\ b_{m+1} & 0 & \cdots & 0 \end{pmatrix}, \quad b_1^T = (b_1, \ldots, b_n) \]

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 \( \alpha \).

If one sets
\[ c_0 = 1, \quad \bar{d}_i = d_i - \alpha, \quad e_0 = 1, \quad b_1 = 0 \quad \text{and} \]
\[ e_i = -\bar{d}_i e_{i-1} + b_i^2 c_{i-1} \]
\[ c_i = c_{i-1} \bar{d}_i \]

then it is easy to show by expanding \( D - \alpha I \) by its \( i \)th row that \( e_i \) contains the \( i \)th principal minor.

2) Deleting a row
\[ A = \begin{pmatrix} \bar{A} \\ a^T \end{pmatrix} \]

If an orthogonal matrix \( P \) is constructed such that \( P \begin{pmatrix} U^T \\ \bar{a} \end{pmatrix} = \begin{pmatrix} \bar{a} \\ 1 \end{pmatrix} \), then because of the orthogonality of the two matrices \( PU^T \) has the form
\[ \begin{pmatrix} U^T & \bar{a} \end{pmatrix} \]
Thus
\[
P \Sigma \Sigma^T P^T = PU^T (\bar{A}^T)(\bar{A}^T)U^T P^T
\]
\[
= \left( \begin{array}{c|c} \bar{U} & 0 \\ \hline 0 & 1 \end{array} \right) \left( \begin{array}{c} \bar{A}^T \bar{A} \\ \bar{A} \end{array} \right) \left( \begin{array}{c} \bar{U}^T \\ 0 \end{array} \right)
\]
\[
= \left( \begin{array}{c|c} \bar{U} \bar{A} \bar{A}^T \bar{U} & \bar{U} \bar{A} \bar{A}^T \bar{U} \\ \hline \bar{A}^T \bar{A}^T \bar{U}^T & \bar{A}^T \bar{A}^T \bar{U}^T \end{array} \right)
\]

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

\[
E^T_m^{-1} P \Sigma \Sigma^T P^T E_m^{-1}.
\]

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

\[
P = R_{m-1}, \ldots, R_1
\]

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

\[
\begin{array}{ccc}
I & c_k & s_k \\
c_k & s_k & 0 \\
s_k & 0 & -c_k \\
\end{array}
\]

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

\[
R_k \cdots R_1 DR^T \cdots R_k^T
\]

and the \( k^{th} \) principal minor of

\[
R_{k-1} \cdots R_1 DR^T \cdots R_{k-1}^T = \prod_{i=1}^{k} d_i
\]
The matrix

\[
R_{k-1} \cdots R_1 DR_1^T R_k^T
\]

has the form

\[
\begin{pmatrix}
X & a \\
 a & d_{k+1} & \cdots & d_n
\end{pmatrix}
\]

If \( R_k \) is applied to its left and right, its \( k \)th principal submatrix would be

\[
\det \begin{pmatrix}
X & c_k^a \\
c_k a & c_k^2 y + s_k^2 d_{k+1}^k
\end{pmatrix} = s_k^2 d_{k+1}^k \det(X) + c_k^2 \prod_{i=1}^k d_i
\]

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

\[
\begin{align*}
  e_0 &= 0, \\
  f_0 &= 1, \\
  f_k &= d_k f_{k-1} \\
  e_k &= s_k^2 d_{k+1}^k e_{k-1} + c_k^2 f_k
\end{align*}
\]

In our case we need the principal minors of

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

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

so that \( d_i = \sigma_i^2 - \alpha \) for \( i \leq r \) and \( d_i = -\alpha \) for \( i > r \).

3) Adding a column

\[
\vec{A} = (a_1 \ A).
\]
Since
\[
\begin{pmatrix}
1 & 0 \\
0 & V^T
\end{pmatrix}
\begin{pmatrix}
\bar{A}^T \\
\bar{A}
\end{pmatrix}
\begin{pmatrix}
\bar{a}^T \\
\bar{a}
\end{pmatrix}
= \begin{pmatrix}
\bar{a}^T \\
\bar{v}^T A T \bar{a} \\
\Sigma T \Sigma
\end{pmatrix}
= D
\]
the singular values of \( \bar{A} \) are the square roots of the eigenvalues of \( D \) and one may proceed as if one were adding a row.

4) Deleting a column
\[
A = (\bar{A} : \bar{a})
\]
If an orthogonal matrix \( P \) is constructed such that
\[
P V^T e_n = e_n',
\]
then
\[
e_n^T P U^T = e_n^T
\]
and because of the orthogonality of the 2 matrices, \( P V^T \) has the form
\[
\begin{pmatrix}
\bar{V} & 0 \\
0 & 1
\end{pmatrix}
\]
Thus
\[
P \Sigma^T \Sigma P = P V^T \begin{pmatrix}
\bar{A}^T \\
\bar{a}
\end{pmatrix}
\begin{pmatrix}
\bar{a}^T \\
\bar{v}^T A T \bar{a} \\
\Sigma T \Sigma
\end{pmatrix}
= \begin{pmatrix}
\bar{V} & 0 \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
\bar{A}^T \\
\bar{a}
\end{pmatrix}
\begin{pmatrix}
\bar{V}^T | 0 \\
0 | 1
\end{pmatrix}
= \begin{pmatrix}
\bar{V} \bar{A}^T \bar{A} V & \bar{V} \bar{A}^T \bar{a} \\
\bar{a}^T \bar{A} V & \bar{a}^T \bar{a}
\end{pmatrix}
\]
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 \( \Lambda = U \Sigma V^T \) and \( \bar{\Lambda} = \begin{pmatrix} \bar{A}^T \\ \bar{a}
\end{pmatrix} \).
Since \( \Lambda^T \Lambda = V \Sigma^T \Sigma V \),
\[
V^T (\bar{a} : \Lambda^T) \begin{pmatrix}
\bar{A}^T \\
\bar{a}
\end{pmatrix} V = \Sigma^T \Sigma + (V^T \bar{a}) (V^T \bar{a})^T
= D + \bar{b} \bar{b}^T
\]
which means that the singular values of \( \tilde{A} \) are the square roots of the eigenvalues of \( D + \tilde{b} \tilde{b}^T \). Bartels, Golub, and Saunders [1] have shown that solving the eigenvalue problem

\[
(D + \tilde{b} \tilde{b}^T) \vec{x} = \lambda \vec{x}
\]

is equivalent to solving the generalized eigenvalue problem

\[(4.2) \quad A \vec{x} = \lambda B \vec{x} \]

where \( A \) and \( B \) are both symmetric and tridiagonal and \( B \) is positive definite and

\[
A = KDK^T + K \tilde{b} \tilde{b}^T K^T \\
B = K K^T
\]

where \( K \) is the bidiagonal matrix such that

\[
K \tilde{b} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b_n \end{pmatrix}
\]


2) Deleting a row.

Let \( A = \begin{pmatrix} \tilde{A} \\ \tilde{\Sigma} \end{pmatrix} \).

Since

\[
\Sigma^T \Sigma = V^T (\tilde{A}^T \tilde{\Sigma}^T \begin{pmatrix} \tilde{A} \\ \tilde{\Sigma} \end{pmatrix} V = V^T \tilde{A} \tilde{\Sigma} V + (V^T \tilde{A}) (V^T \tilde{\Sigma})^T
\]

the singular values of \( \tilde{A} \) are the square roots of the eigenvalues of \( D - \tilde{b} \tilde{b}^T \)

where \( D = \Sigma^T \Sigma \) and \( \tilde{b} = V \tilde{\Sigma} \). As in the previous case these
eigenvalues can be found by solving a generalized eigenvalue problem like (4.2). In this case

\begin{equation}
A = KDK^T - (K \tilde{b}) (K \tilde{b})^T.
\end{equation}

3) Adding a column.

Let $\tilde{A} = (\tilde{a}_2; A)$.

Since

$$U^T(\tilde{a}; A)(\tilde{a}_2^T)U = (U^T\tilde{a})(U^T\tilde{a}_2)^T + \Sigma \Sigma^T,$$

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

4) Deleting a column.

Let $A = (a; \tilde{A})$.

Since

$$\Sigma \Sigma^T = U^T(\tilde{a}; A)(\tilde{a}_2^T)U = U^TAA^TU + U^T\tilde{a}_2\tilde{a}_2^TU$$

the singular values of $\tilde{A}$ are the square roots of the nonzero eigenvalues of (4.3) with $D = \Sigma \Sigma^T$ and $\tilde{b} = U^T\tilde{a}_2$.

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


/EY