Numer. Math. 31, 111–129 (1978)

Numerische Mathematik

Updating the Singular Value Decomposition*

James R. Bunch and Christopher P. Nielsen

Department of Mathematics, University of California, San Diego, La Jolla, CA92093, USA

Summary. Let A be an $m\times n$ matrix with known singular value decomposition. The computation of the singular value decomposition of a matrix $\tilde{A}$ is considered, where $\tilde{A}$ is obtained by appending a row or a column to A when $m\geqq n$ or by deleting a row or a column from A when $m>n$. An algorithm is also presented for solving the updated least squares problem $\tilde{A}y-\tilde{b}$, obtained from the least squares problem $Ax-b$ by appending an equation, deleting an equation, appending an unknown, or deleting an unknown.

Subject Classifications. AMS (MOS): 65F30; CR: 5.14.

1. Introduction

The singular value decomposition is of great importance for calculating the rank of a matrix and for obtaining least squares solutions of overdetermined systems of linear equations [7]. Let $\mathbb{R}$ denote the real numbers and $\mathbb{R}^{m\times n}$ denote the set of $m\times n$ real matrices. Let $A\in\mathbb{R}^{m\times n}$, where $m\geqq n$. There exist matrices U, Σ, and V such that $A=U\Sigma V^T$, the *singular value decomposition* (SVD) of A; $U\in\mathbb{R}^{m\times n}$ has orthonormal columns, $V\in\mathbb{R}^{n\times n}$ is orthogonal, and $\Sigma=\operatorname{diag}(\sigma_1,\dots,\sigma_r,0,\dots,0)\in R^{n\times n}$, $\sigma_1\geqq\sigma_2\geqq\cdots\geqq\sigma_r>0$, where r is the rank of A.

We say that x is a least squares solution of $Ay-b$ if $\|Ax-b\|_2=\min_y\|Ay-b\|_2$. If $r=n$, then there is a unique least squares solution x, but if $r<n$, then there are infinitely many such x. However, there is a unique least squares solution of minimum 2-norm, i.e. there exists a unique $\hat{x}$ such that $\|\hat{x}\|_2=\min_{x\in S}\|x\|_2$ where $S=\{x\colon\|Ax-b\|_2=\min_y\|Ay-b\|_2\}$. Then $\hat{x}$ is given by $\hat{x}=V\Sigma^+U^Tb$ where $\Sigma^+=\operatorname{diag}(\sigma_1^{-1},\sigma_2^{-1},\dots,\sigma_r^{-1},0,\dots,0)$. For further discussion, see [3–7, 9].

Having calculated the singular value decomposition of A or having obtained a least squares solution to $Ax-b$, we often need to modify A by appending a

* This research was supported by NSF grants MCS 75-06510 and MCS 76-03139

0029-599X/78/0031/0111/$03.80

row or column or by deleting a row or column [2–6]. We would like to take advantage of the singular value decomposition of A in order to reduce the amount of work necessary when computing the singular value decomposition of the modified matrix and when obtaining the new least squares solution.

These are called updating problems and can be stated formally as:

I. given the SVD of an $m\times n$ matrix $A=U\Sigma V^T$, make use of U, Σ, and V in the computation of the SVD of $\tilde{A}=\tilde{U}\tilde{\Sigma}\tilde{V}^T$, where

(a) $\tilde{A}=\left[\begin{array}{c}A\\ \hdashline a^T\end{array}\right]$, i.e. appending a row, $m\geqq n$;

(b) $\tilde{A}=[A\,\vdots\,a]$, i.e. appending a column, $m\geqq n$;

(c) $A=\left[\begin{array}{c}\tilde{A}\\ \hdashline a^T\end{array}\right]$, i.e. deleting a row, $m>n$;

(d) $A=[\tilde{A}\,\vdots\,a]$, i.e. deleting a column, $m>n$;

II. given Σ, V, and $c=U^Tb$, compute $\tilde{\Sigma}$, $\tilde{V}$ and an updated right-hand side $\tilde{c}$ $=\tilde{U}^T\left[\begin{array}{c}b\\ \hdashline b_{m+1}\end{array}\right]=\tilde{U}^T\tilde{b}$.

Businger [2] considers problem II and presents an algorithm which requires $2n^3+O(n^2)$ multiplications for the direct part of the algorithm and approximately $4n^3+O(n^2)$ multiplications for the iterative part.

In Sections 2–5 we use the results in [1] to present algorithms for solving problems $I(a)$–(d), respectively, in which both the left and right singular vectors are explicitly constructed once the new singular values have been calculated by solving a rank-one updated eigenvalue problem. Our algorithms require $n^3+mn^2+7n^2+O(n)$ multiplications for the direct part of each algorithm and approximately $12n^2+O(n)$ multiplications for the iterative part. Computing the SVD of an $m\times n$ matrix requires $n^3+mn^2+O(mn)$ multiplications for the direct part and $kmn^2+kn^3+O(mn)$ multiplications for the iterative part, if k iterations (on average) are needed to find each singular value.

In Section 6 we use the results of Sections 2–5 and present an algorithm to solve the least squares problem II, where, again, the right singular vectors and the updated right-hand side can be calculated directly, once the new singular values have been calculated. Our algorithm requires $n^3+6n^2+O(n)$ multiplications for the direct part and $12n^2+O(n)$ for the iterative part. The amount of work is independent of m so our algorithm is of interest primarily when $m\gg n$. (This situation is common in statistical applications.)

Numerical results are given in Section 7.

All results follow analogously for complex matrices.

2. Updating the Singular Value Decomposition when Appending a Row

Let $A\in\mathbb{R}^{m\times n}$, $m\geqq n$, have the SVD decomposition

$$\begin{aligned} A=U\Sigma V^T, \quad & U\in\mathbb{R}^{m\times n}, \quad U^TU=I,\\ & \Sigma\in\mathbb{R}^{n\times n}, \\ & V\in\mathbb{R}^{n\times n}, \quad V^TV=VV^T=I, \end{aligned} \tag{1}$$

and consider the decomposition

$$\begin{aligned} A=\hat{U}\hat{\Sigma}V^T, \quad & \hat{U}\in\mathbb{R}^{m\times m}, \quad \hat{U}^T\hat{U}=\hat{U}\hat{U}^T=I, \\ & \hat{\Sigma}\in\mathbb{R}^{m\times n}, \\ & V\in\mathbb{R}^{n\times n}, \quad VV^T=V^TV=I, \end{aligned} \tag{2}$$

where $\hat{\Sigma}=\begin{bmatrix}\Sigma\\0\end{bmatrix}$, $\hat{U}=(U, U_2)$, and Σ, U are given in (1); U_2 is any $m\times(m-n)$ matrix satisfying $U_2^T U_2=I$ and $U^T U_2=0$.

Let $\tilde{A}=\begin{pmatrix}A\\a^T\end{pmatrix}$ and $\tilde{A}=\tilde{U}\tilde{\Sigma}\tilde{V}^T$, where

$$\begin{aligned} &\tilde{U}\in\mathbb{R}^{(m+1)\times n}, \quad \tilde{U}^T\tilde{U}=I, \\ &\tilde{\Sigma}\in\mathbb{R}^{n\times n}, \\ &\tilde{V}\in\mathbb{R}^{n\times n}, \quad \tilde{V}^T\tilde{V}=I=\tilde{V}\tilde{V}^T. \end{aligned}$$

If $a=0$, then $\tilde{U}=\begin{bmatrix}U\\ \hline 0^T\end{bmatrix}$, $\tilde{\Sigma}=\Sigma$, and $\tilde{V}=V$. Let us assume $a\neq 0$.

2.1. Rank-one Modification of the Symmetric Eigenproblem

Since $\tilde{A}^T\tilde{A}=A^TA+aa^T=V\Sigma^2V^T+aa^T=\tilde{V}\tilde{\Sigma}^2\tilde{V}^T$, we can compute $\tilde{V}$ and $\tilde{\Sigma}$ from the solution of the following eigenvalue problem: given a symmetric matrix B with known eigensystem $B=QDQ^T$, $Q^TQ=QQ^T=I$, calculate the eigensystem of $\tilde{B}=B+\rho bb^T$, $b^Tb=1$, $\rho\in R$.

This problem has been considered in [1]. We shall describe only the real case, but all results hold analogously for the rank-one modification of Hermitian eigenproblems.

The algorithm is based on the following

Theorem 1. Let $C=D+\rho zz^T$, where D is diagonal, $\|z\|_2=1$. Let $d_1\leqq d_2\leqq\cdots\leqq d_n$ be the eigenvalues of D, and let $\tilde{d}_1\leqq\tilde{d}_2\leqq\cdots\leqq\tilde{d}_n$ be the eigenvalues of C. Then $\tilde{d}_i=d_i+\rho\mu_i$, $1\leqq i\leqq n$, where $\sum_{i=1}^{n}\mu_i=1$, and $0\leqq\mu_i\leqq 1$. Moreover, $d_1\leqq\tilde{d}_1\leqq d_2\leqq\cdots\leqq d_n\leqq\tilde{d}_n$ if $\rho>0$ and $\tilde{d}_1\leqq d_1\leqq\tilde{d}_2\leqq\cdots\leqq\tilde{d}_n\leqq d_n$ if $\rho<0$. Finally, if the d_i are distinct and all the elements of z are nonzero, then the eigenvalues of C strictly separate those of D.

A version of the above theorem may be found in [8] and in [9, pp. 95–98].

The algorithm proceeds in three stages:

(i) initial deflation;
(ii) solution of a nonlinear equation for the eigenvalues of $\tilde{B}$;
(iii) explicit construction of the eigenvectors of $\tilde{B}$.

(i) Deflation

The given problem may be simplified at the outset by making a few observations. First of all, note that

$$B+\rho b b^T = Q(D+\rho z z^T)Q^T,$$

where $b=Qz$. Thus, if $D+\rho z z^T = X\tilde{D}X^T$ is the orthogonal decomposition of $D+\rho z z^T$ then the orthogonal decomposition of $\tilde{B}$ is

$$\tilde{B}=\tilde{Q}\tilde{D}\tilde{Q}^T,$$

where $\tilde{Q}=QX$. Without loss of generality, we shall assume throughout that $\|z\|_2=1$. Let $z=[\zeta_1, \dots, \zeta_n]^T$.

There are some situations for which we can deflate the problem immediately.

If $\zeta_i=0$ for some i, then $\tilde{d}_i=d_i$ for this i and the corresponding eigenvector remains unchanged, since $(D+\rho z z^T)e_i = De_i+\rho z \zeta_i = d_i e_i$ as $\zeta_i=0$.

If $|\zeta_i|=1$ for some i (so $\zeta_j=0$, $j\neq i$), then $\tilde{d}_i=d_i+\rho$, $\tilde{d}_j=d_j$, $j\neq i$, and all eigenvectors remain unchanged.

If B has multiple eigenvalues, then let us suppose that λ is an eigenvalue of $B=QDQ^T$ of multiplicity $r\geqq 2$, and let $Q_1\in\mathbb{R}^{n\times r}$ span the subspace associated with λ. Partition Q and permute D so that $Q=(Q_1, Q_2)$. Note that all we require of Q_1 is that $Q_1^TQ_1=I$ and $Q_1^TQ_2=0$. We now have

$$z=Q^Tb=\begin{bmatrix} Q_1^T b \\ Q_2^T b \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

Let $H\in\mathbb{R}^{r\times r}$ be an elementary reflector such that $Hz_1=-\sigma e_1$, $\sigma=\|z_1\|_2$. Define $\bar{Q}_1\equiv Q_1H^T$. Then

(a) $\bar{Q}_1^T b = HQ_1^T b = Hz_1 = -\sigma e_1$,

(b) $\bar{Q}_1^T\bar{Q}_1 = HQ_1^TQ_1H^T = HH^T = I$,

(c) $\bar{Q}_1^TQ_2 = HQ_1^TQ_2 = 0$.

Thus we may replace Q_1 by $\bar{Q}_1$ and hence let $Q=(\bar{Q}_1, Q_2)$. This has the advantage of introducing $r-1$ zero components into z, and hence we will have $r-1$ eigenvalues and eigenvectors which will not change. We repeat the above procedure for all sets of multiple eigenvalues.

After performing this procedure, permute the columns of Q and the diagonal elements of D to correspond to the zero components of z, i.e.

$$Q=(Q_1, Q_2), \quad D=\begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}, \quad z=\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}=\begin{bmatrix} 0 \\ z_2 \end{bmatrix}.$$

The problem is then to update

$$D+\rho z z^T = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} + \rho \begin{bmatrix} 0 & 0 \\ 0 & z_2 z_2^T \end{bmatrix};$$

that is, we compute the eigensystem of $(D_2+\rho z_2 z_2^T)\in\mathbb{R}^{k\times k}$, $k=n-r+1$. If $\bar{Q}_2$ is the matrix whose columns are the eigenvectors of $D_2+\rho z_2 z_2^T$, then the matrix of eigenvectors of $\tilde{B}=B+\rho b b^T$ is given by $\tilde{Q}=(Q_1, Q_2\bar{Q}_2)$.

(ii) The Eigenvalues of $D+\rho z z^T$

Let us assume that we are working with an $n\times n$ problem for which no deflation is possible. Thus we consider the problem of finding the updated eigensystem of $C\equiv D+\rho z z^T$, where the d_i are distinct and $\zeta_i\neq 0$ for all i. Golub [3] has shown that in the above situation the eigenvalues of C are the zeros of $w(\lambda)$, where

$$w(\lambda)=1+\rho\sum_{j=1}^{n}\frac{\zeta_j^2}{(d_j-\lambda)}.$$

Let us denote the eigenvalues of C by $\tilde{d}_1<\tilde{d}_2<\cdots<\tilde{d}_n$. Since $\tilde{d}_i=d_i+\rho\,\mu_i$ satisfies $w(\tilde{d}_i)=0$, we can also compute the eigenvalues of C by solving $w_i(\mu)=0$, where

$$w_i(\mu)=1+\sum_{j=1}^{n}\frac{\zeta_j^2}{(\delta_j^i-\mu)}$$

and $\delta_j^i=(d_j-d_i)/\rho$. This reformulation has some numerical advantages with respect to the accurate determination of the updated eigenvectors.

We turn now to the numerical solution of $w_i(\mu)=0$. We shall assume that $\rho>0$ and $1\leqq i<n$. There is no loss of generality in assuming that $\rho>0$; otherwise, we can replace d_i by $-d_{n-i+1}$ and ρ by $-\rho$. The case $i=n$ is special and we shall deal with it separately. Let

$$\psi(t)\equiv\sum_{j=1}^{i}\frac{\zeta_j^2}{(\delta_j^i-t)}\quad\text{and}\quad\varphi(t)\equiv\sum_{j=i+1}^{n}\frac{\zeta_j^2}{(\delta_j^i-t)}.$$

Then we are seeking $0<\mu_i<\min\left(1-\sum_{j=1}^{i-1}\mu_j,\ \delta_{i+1}^i\right)$ such that

$$-\psi(\mu_i)=\varphi(\mu_i)+1.$$

Observe that the terms in the sums that define φ and ψ are all of the same sign. Moreover, $1+\varphi(t)$ is an increasing convex function and $-\psi(t)$ is a decreasing convex function of t on the interval $(0,\delta_{i+1}^i)$.

Newton's method, safeguarded with bisection, would be an obvious choice for finding the solutions to $w_i(\lambda)=0$. However, Newton's method is based on a local linear approximation to the function w_i. Since our functions φ and ψ are rational functions, it seems more natural to develop a method based on a local approximation to φ and ψ via simple rational functions. For the moment we shall consider a local approximation to φ and ψ at a point t_1 such that $0<t_1<\mu_i$. Our goal is to obtain a new approximation t_2 to the root μ_i. Hence we define the interpolating rationals $\frac{p}{q-t}$, $r+\frac{s}{\delta-t}$, (where $\delta=\delta_{i+1}^i$) such that

$$\frac{p}{q-t_1}=\psi_1,\qquad r+\frac{s}{\delta-t_1}=\varphi_1,$$

$$\frac{p}{(q-t_1)^2}=\psi_1',\qquad \frac{s}{(\delta-t_1)^2}=\varphi_1',$$

where $\psi_1=\psi(t_1)$, $\psi_1'=\psi'(t_1)$, etc. It is easily verified that

$$p=\psi_1^2/\psi_1',\qquad r=\varphi_1-\Delta\,\varphi_1',$$

$$q=t_1+\psi_1/\psi_1',\qquad s=\Delta^2\,\varphi_1',$$

where $\Delta=\delta-t_1$.

Our next approximation t_2 to μ_i is obtained by solving the equation

$$\frac{-p}{q-t_2}=1+r+\frac{s}{\delta-t_2}.$$

This only requires the solution of a quadratic equation. The ambiguity of sign is easily dispensed with and we find the correct formula to be

$$t_2=t_1+2b/(a+\sqrt{a^2-4b}),$$

where

$$a=(\Delta(1+\varphi_1)+\psi_1^2/\psi_1')/c+\psi_1/\psi_1',$$

$$b=(\Delta w\psi_1)/(\psi_1' c),$$

$$c=1+\varphi_1-\Delta\varphi_1',$$

and

$$w=1+\varphi_1+\psi_1.$$

Moreover, given $t_0\in(0,\mu_i)$ we have shown in [1] that a sequence $\{t_k\}$, such that t_{k+1} is obtained from t_k as t_2 is obtained from t_1, has a quadratic rate of convergence. Since the previous discussion implies that the convergence is monotone we have, in theory, avoided the need to safeguard the method.

The case $i=n$ is special since $\varphi\equiv 0$. Thus, when $i=n$ we want to solve $-\psi(t)=1$; the iteration is now given by

$$t_{k+1}=t_k+\left(\frac{1+\psi_k}{\psi_k'}\right)\psi_k.$$

Note that $t_0<t_k<\mu_n$ implies $-\psi_k\downarrow 1$ as $k\to+\infty$; thus the iteration may be viewed as an accelerated Newton's method for finding a zero of $1+\psi(t)$.

(iii) Calculating the Eigenvectors

After we have calculated the eigenvalues $\tilde{d}_1,\ldots,\tilde{d}_n$ of $\tilde{B}=B+\rho bb^T$, we need to solve

$$\tilde{B}\tilde{q}_i-\tilde{d}_i\tilde{q}_i=0,\qquad 1\leqq i\leqq n,$$

for the eigenvectors $\tilde{q}_i$ of $\tilde{B}$. If $z=Q^Tb$, and $x_i=Q^T\tilde{q}_i$, then multiplying on the left by Q^T gives

$$(D_i+\rho zz^T)x_i=0,$$

where $D_i=D-\tilde{d}_iI$. Since we are working with deflated systems, Theorem 1 implies that D_i is nonsingular.

We may calculate $\tilde{q}_i$ for $1\leqq i\leqq n$ by the formula

$$\tilde{q}_i=\frac{QD_i^{-1}z}{\|D_i^{-1}z\|_2}.$$

In practice, however, we do not recommend forming $D_i^{-1}z$ explicitly, but instead we compute it by the zero finding algorithm. Since we are solving $0 = w_i(\mu) = 1 + \sum_{j=1}^{n} \frac{\zeta_j^2}{(\delta_j^i - \mu)}$ for μ_i, we note that if $\mathscr{D}_i \equiv \frac{1}{\rho} D_i$, then $\mathscr{D}_i = \mathrm{diag}(\Delta_1^i, \Delta_2^i, \ldots, \Delta_n^i)$, where $\delta_j^i = \frac{(d_j - d_i)}{\rho}$ and $\Delta_j^i = \delta_j^i - \mu_i$, $1 \leqq i, j \leqq n$.

The quantities Δ_j^i may be maintained by the zero finding algorithm as follows: consider computing the i^{th} updated eigenvalue; obtain the initial guess $\mu_i^{(0)}$; initialize $\Delta_j^i = \delta_j^i - \mu_i^{(0)}$, $1 \leqq j \leqq n$. Each time a new iterate $\mu_i^{(k)}$ is computed, set $\Delta_j^i \leftarrow \Delta_j^i + (\mu_i^{(k-1)} - \mu_i^{(k)})$. Note that the corrections are all of the same sign.

The advantage of this method is that for roots which are poorly separated, these quantities do not suffer the loss of accuracy due to cancellation that will occur when $d_i - \tilde{d}_i$ is formed. Instead, the iterative corrections to the root μ_i are applied directly to the Δ_j^i; these corrections are likely to be quite small compared to Δ_j^i, even when d_i and d_{i+1} are close. Since $\mathscr{D}_i^{-1} z = \rho D_i^{-1} z$ have the same direction, we may use the Δ_j^i directly to compute the updated eigenvector. For further details on this point, see [1].

The computation of each eigenvector requires $n^2 + O(n)$ multiplications; so the computation of $\tilde{Q}$ requires $n^3 + O(n^2)$ multiplications. Note that if we have deflated to a $k \times k$ eigenproblem, then the operation count is $nk^2 + O(k^2)$.

(iv) Sensitivity of the Eigenvectors

Let $C = D + \rho z z^T$, and consider the i^{th} eigenvector x_i of C corresponding to eigenvalue $\tilde{d}_i$. We assume as before that we are working with a deflated system, i.e. all the d_j are distinct and all $\zeta_j \neq 0$. Let $\theta_i = \min_j |d_j - \tilde{d}_i|$. Suppose we compute an approximate eigenvalue $\hat{d}_i$ using the nonlinear equation solver. Let $\hat{d}_i = \tilde{d}_i + \varepsilon$, and assume $|\varepsilon| < \beta \theta_i$ where $0 < \beta < 1$. The exact eigenvector x_i is given by

$$x_i = \frac{D_i^{-1} z}{\|D_i^{-1} z\|_2}, \qquad D_i = D - \tilde{d}_i I.$$

Then, the corresponding eigenvector $\tilde{q}_i$ of $\tilde{B}$ is given by

$$\tilde{q}_i = Q x_i.$$

However, if $\hat{d}_i$ is used in place of $\tilde{d}_i$ above, we obtain the corresponding approximate eigenvector of C as

$$\hat{x} = \frac{\hat{D}_i^{-1} z}{\|\hat{D}_i^{-1} z\|_2}, \qquad \hat{D}_i = D - \hat{d}_i I;$$

the corresponding approximate eigenvector for $\tilde{B}$ is

$$\hat{q}_i = Q \hat{x}_i.$$

In [1], it is shown that

$$\|\tilde{q}_i - \hat{q}_i\|_2 = \|x_i - \hat{x}_i\|_2 \leqq \frac{2\sqrt{3}\,\varepsilon}{(1-\beta)\,\theta_i}.$$

2.2. Calculating $\tilde{\Sigma}$, $\tilde{V}$, and $\tilde{U}$

We may assume that our SVD problem has already been deflated, i.e. we may assume $\sigma_1 > \cdots > \sigma_n \geqq 0$. Since $a \neq 0$, we have $\rho > 0$. If we assume that the components of $V^T a$ are nonzero, then $\tilde{\sigma}_1 > \sigma_1 > \tilde{\sigma}_2 > \cdots > \tilde{\sigma}_n > \sigma_n \geqq 0$, and hence the $\tilde{\sigma}_i$ are all positive, $1 \leqq i \leqq n$.

In Section 2.1, let $B = A^T A$, $Q = V$, $D = \Sigma^2$, $\alpha = 1/\|a\|_2$, $b = \alpha a$, and $\rho = (1/\alpha^2) > 0$. Thus, if $\tilde{B} = \tilde{Q}\tilde{D}\tilde{Q}^T$, we may take $\tilde{V} = \tilde{Q}$ and $\tilde{\sigma}_i = \tilde{d}_i^{\frac{1}{2}}$ for $1 \leqq i \leqq n$, where $\tilde{D} = \operatorname{diag}(\tilde{d}_1, \ldots, \tilde{d}_n)$. (Note that $d_1 \geqq d_2 \geqq \cdots \geqq d_n \geqq 0$, where $D = \operatorname{diag}(d_1, \ldots, d_n)$, and $\rho > 0$ imply that $\tilde{d}_i > 0$ for $1 \leqq i \leqq n$.)

The updated right singular vectors $\tilde{V} = (\tilde{v}_1, \ldots, \tilde{v}_n)$ are given by

$$\tilde{v}_i = \beta_i V T_i^{-1} z,$$

where

$$\beta_i = \frac{1}{\|T_i^{-1} z\|_2}, \qquad z = \alpha V^T a,$$

$$\alpha = \frac{1}{\|a\|_2} = \frac{1}{\|V^T a\|_2}, \quad \text{and} \quad T_i = \Sigma^2 - \tilde{\sigma}_i^2 I.$$

(Since $\tilde{\sigma}_i \neq \sigma_j$ for $1 \leqq j \leqq n$, T_i is nonsingular.)

We may compute $\tilde{V}$ in the following order; we shall count divisions as multiplications.

1) $\alpha \leftarrow 1/\|a\|_2$, $n+1$ multiplications;
2) $y \leftarrow V^T a$, n^2 multiplications;
3) $z \leftarrow \alpha y$, n multiplications;
4) for each i, $1 \leqq i \leqq n$:
   (a) $T_i^{-1} \leftarrow (\Sigma^2 - \tilde{\sigma}_i^2 I)^{-1}$, $3n$ multiplications;
   (b) $x_i \leftarrow T_i^{-1} z$, n multiplications;
   (c) $\beta_i \leftarrow 1/\|x_i\|_2$, $n+1$ multiplications;
   (d) $w_i \leftarrow V x_i$, n^2 multiplications;
   (e) $\tilde{v}_i \leftarrow \beta_i w_i$, n multiplications.

Hence, we may compute V with $n^3 + 7n^2 + 3n + 1$ multiplications.

Since $\tilde{A} = \tilde{U}\tilde{\Sigma}\tilde{V}^T$ implies $\tilde{A}\tilde{V} = \tilde{U}\tilde{\Sigma}$, the updated left singular vectors $\tilde{U} = (\tilde{u}_1, \ldots, \tilde{u}_n)$ are related to the updated right singular vectors by

$$\tilde{A}\tilde{v}_i = \tilde{\sigma}_i \tilde{u}_i \quad \text{for } 1 \leqq i \leqq n.$$

Thus, if the matrix A (or $\tilde{A}$) is still available, we can compute the $\tilde{u}_i$ by:

$$\tilde{u}_i = \frac{1}{\tilde{\sigma}_i} \tilde{A} \tilde{v}_i = \frac{1}{\tilde{\sigma}_i} \begin{bmatrix} A \\ a^T \end{bmatrix} \tilde{v}_i, \qquad 1 \leqq i \leqq n.$$

This requires $(m+1)(n+1)$ multiplications for each $\tilde{u}_i$, and hence $mn^2 + mn + n^2 + n$ multiplications for $\tilde{U}$.

If the matrix A is no longer available, we must compute the $\tilde{u}_i$ by using U and Σ. But, for $1 \leqq i \leqq n$,

$$\tilde{u}_i=\frac{1}{\tilde{\sigma}_i}\begin{bmatrix}A\\ \hdashline a^T\end{bmatrix}\tilde{v}_i=\frac{\beta_i}{\tilde{\sigma}_i}\begin{bmatrix}U\Sigma V^T\\ \hdashline a^T\end{bmatrix}VT_i^{-1}z=\frac{\beta_i}{\tilde{\sigma}_i}\begin{bmatrix}U\Sigma T_i^{-1}z\\ \hdashline -\alpha\end{bmatrix},$$

since $z^T T_i^{-1} z=\sum_{j=1}^{n}\frac{\zeta_j^2}{(\sigma_j^2-\tilde{\sigma}_i^2)}=-\alpha^2$, where β_i, z, α, and T_i are as above. If α, β_i, and $x_i=T_i^{-1}z$ were already computed and saved above, then forming $\tilde{U}$ would require $mn^2+n^2+2mn+2n$ multiplications.

Thus, computing $\tilde{U}$ and $\tilde{V}$ requires $mn^2+n^3+8n^2+mn+3n+1$ multiplications if the matrix A is still available and $mn^2+n^3+8n^2+2mn+5n+1$ multiplications if A is no longer available.

Using the sensitivity results for the symmetric eigenproblem in Section 2.1, we have that if we have computed an approximate eigenvalue $\hat{d}_i$ such that $\hat{d}_i=\tilde{d}_i+\varepsilon$, where $|\varepsilon|<\beta\theta_i$, $0<\beta<1$, $\theta_i=\min_j|d_j-\tilde{d}_i|$, then

$$\tilde{\sigma}_i=\tilde{d}_i^{\frac{1}{2}} \quad\text{and}\quad \hat{\sigma}_i\equiv\hat{d}_i^{\frac{1}{2}};$$

if the resulting $\tilde{v}_i$, $\tilde{u}_i$ are perturbed to $\hat{v}_i$, $\hat{u}_i$, respectively, then $\|\tilde{v}_i-\hat{v}_i\|_2$ and $\|\tilde{u}_i-\hat{u}_i\|_2$ are bounded by

$$\frac{2\sqrt{3}\,|\hat{d}_i-\tilde{d}_i|}{(1-\beta)\,\theta_i}=\frac{2\sqrt{3}\,|\hat{\sigma}_i^2-\tilde{\sigma}_i^2|}{(1-\beta)\min_j|\sigma_j^2-\tilde{\sigma}_i^2|}.$$

2.3. Another Approach

Businger's SVD update when appending a row can be stated as follows.

Consider $\tilde{A}=\begin{bmatrix}A\\ a^T\end{bmatrix}$. Perform an orthogonal equivalence transformation to obtain $\begin{bmatrix}\Sigma\\ \alpha z^T\end{bmatrix}$, and bidiagonalize to obtain $\begin{bmatrix}B\\ 0\end{bmatrix}$. Then perform the QR-variant algorithm to obtain $\begin{bmatrix}\tilde{\Sigma}\\ 0\end{bmatrix}$. Each phase produces a sequence of Givens rotations which are applied to V and U to obtain $\tilde{V}$ and $\tilde{U}$, or, if we are updating least squares solutions, to V and c to obtain $\tilde{V}$ and $\tilde{c}$. For least squares problems, bidiagonalization requires $2n^3+O(n^2)$ multiplications, and the QR phase requires $\approx 4n^3$ multiplications. Thus the bulk of the work is in applying rotations to V during the QR phase.

We could combine our method with Businger's method by performing the bidiagonalization and QR iteration on $\begin{bmatrix}\Sigma\\ z^T\end{bmatrix}$. The rotations would not be applied to V or c; rather, we use our method of calculating $\tilde{V}$ and $\tilde{c}$ given the updated singular values. However, this method involves forming $T_i^{-1}z$ explicitly, which can cause problems with stability. This "hybrid" algorithm requires $n^3+O(n^2)$ multiplications for updating least squares problems.

Similarly, all four updates (and their corresponding least squares problems) could be expressed in terms of a "hybrid" algorithm consisting of bidiagonalization, a QR phase, and singular vector updates.

3. Updating the SVD when Deleting a Row

By permuting rows we may assume, without loss of generality, that the last row is to be deleted from a matrix whose SVD is known.

Let $A=\begin{bmatrix}\tilde{A}\\ a^T\end{bmatrix}$; a^T is the row to be deleted. Let $A=U\Sigma V^T$, $U^TU=I$, and $V^TV=I$. We wish to calculate the SVD of $\tilde{A}$, given the SVD of A; assume $a\neq 0$.

3.1. Calculating $\tilde{V}$ and $\tilde{\Sigma}$

Since $\tilde{A}^T\tilde{A}=A^TA-aa^T=V\Sigma^2V^T-aa^T$, we may calculate the new right singular vectors $\tilde{V}$ and the new singular values $\tilde{\Sigma}$ via the eigensystem update [1].

In order to deflate the problem, we will assume that all singular values are distinct and that all components of $z=\dfrac{V^Ta}{\|V^Ta\|_2}$, where $z=(\zeta_1,\ldots,\zeta_n)$, are non-zero. In addition, zero singular values are automatically deflated out by the following

Theorem 2. If $\sigma_k=0$, then $\zeta_k=0$, where $z=\alpha V^Ta$, $\alpha=1/\|a\|_2$, and $A=\begin{bmatrix}\tilde{A}\\ a^T\end{bmatrix}=U\Sigma V^T$.

Proof. We have $A=\begin{bmatrix}\tilde{A}\\ a^T\end{bmatrix}=\begin{bmatrix}U_{m-1}\\ u^T\end{bmatrix}\Sigma V^T$, where $U=\begin{bmatrix}U_{m-1}\\ u^T\end{bmatrix}$ is partitioned in the same way that A is.

Then $a^T=u^T\Sigma V^T$, or $\alpha z=V^Ta=\Sigma u$. Thus, if $\sigma_k=0$, we must have $\zeta_k=0$. QED.

Thus we may assume $\sigma_1>\sigma_2>\cdots>\sigma_n>0$ and $\zeta_k\neq 0$ for $1\leqq k\leqq n$.

After we have deflated the problem and calculated the new singular values, the updated right singular vectors $\tilde{v}_i$ are given by

$$\tilde{v}_i=\beta_i VT_i^{-1}z, \quad \text{where} \quad \beta_i=\frac{1}{\|T_i^{-1}z\|_2}, \qquad T_i=\Sigma^2-\tilde{\sigma}_i^2 I,$$

and $\tilde{\sigma}_i$ is the i^{th} updated singular value. The cost for obtaining $\tilde{V}$ and $\tilde{\Sigma}$ is $n^3+O(n^2)$ multiplications.

3.2. Calculating $\tilde{U}$

If the matrix $\tilde{A}$ (or A) is still available, we can calculate $\tilde{u}_i=\dfrac{1}{\tilde{\sigma}_i}\tilde{A}\tilde{v}_i$ for $1\leqq i\leqq n$. This requires $(m-1)(n+1)$ multiplications for each $\tilde{u}_i$, and hence $\tilde{U}$ can be computed with mn^2+mn-n^2-n multiplications.

If the matrix $\tilde{A}$ (or A) is not available, let $U=\begin{bmatrix}U_{m-1}\\ \hdashline u^T\end{bmatrix}$, i.e. let U_{m-1} be the first $m-1$ rows of U. Since

$$A=\begin{bmatrix}\tilde{A}\\ \hdashline a^T\end{bmatrix}=U\Sigma V^T=\begin{bmatrix}U_{m-1}\\ \hdashline u^T\end{bmatrix}\Sigma V^T=\begin{bmatrix}U_{m-1}\Sigma V^T\\ \hdashline u^T\Sigma V^T\end{bmatrix},$$

we have

$$\tilde{u}_i = (1/\tilde{\sigma}_i)\,\tilde{A}\,\tilde{v}_i = (1/\tilde{\sigma}_i)\,U_{m-1}\Sigma V^T \tilde{v}_i$$
$$= (\beta_i/\tilde{\sigma}_i)\,U_{m-1}\Sigma T_i^{-1} z,$$

where

$$\beta_i = \frac{1}{\|T_i^{-1} z\|_2}, \qquad z = \alpha V^T a, \qquad \alpha = \frac{1}{\|a\|_2}, \quad \text{and} \quad T_i = \Sigma^2 - \tilde{\sigma}_i^2 I.$$

Calculating $\tilde{U}$ requires $mn^2 + O(n^2)$ multiplications. Thus, updating the SVD of A when deleting a row can be done with $mn^2 + n^3 + O(n^2)$ multiplications.

4. Updating the SVD when Appending a Column

Suppose we know the SVD of A, and we wish to calculate the SVD of $\tilde{A} = (A, a)$. As usual, we let

$$A = U\Sigma V^T, \qquad U^T U = I, \qquad V^T V = I = V V^T,$$
$$A = \hat{U}\hat{\Sigma}V^T, \qquad \hat{U}\hat{U}^T = \hat{U}^T\hat{U} = I,$$
$$\hat{\Sigma} = \begin{bmatrix}\Sigma\\ 0\end{bmatrix}, \qquad \hat{U} = (U, U_2),$$

$A \in \mathbb{R}^{m\times n}$, $U \in \mathbb{R}^{m\times n}$, $V \in \mathbb{R}^{n\times n}$, $\Sigma \in \mathbb{R}^{n\times n}$, $\hat{U} \in \mathbb{R}^{m\times m}$, and $\hat{\Sigma} \in \mathbb{R}^{m\times n}$.

4.1. Calculating $\tilde{\Sigma}$

Consider $\tilde{A}\tilde{A}^T = AA^T + aa^T$. Then $\hat{U}^T\tilde{A}\tilde{A}^T\hat{U} = B = \hat{\Sigma}\hat{\Sigma}^T + (1/\alpha^2)\,\hat{w}\,\hat{w}^T$, where

$$\hat{w} = \alpha\hat{U}^T a = \alpha\begin{bmatrix}U^T a\\ U_2^T a\end{bmatrix} = \begin{bmatrix}\bar{w}\\ \bar{w}_2\end{bmatrix} \quad \text{and} \quad \alpha = 1/\|a\|_2.$$

The nonzero eigenvalues of $\tilde{A}\tilde{A}^T$ (at most $n+1$) can be found by solving the following secular equation:

$$f(\lambda) = 1 + (1/\alpha^2)\sum_{j=1}^{n}\frac{\bar{w}_j^2}{(\sigma_j^2 - \lambda)} - \frac{\bar{\tau}^2}{\alpha^2\lambda},$$
$$\bar{\tau}^2 = \|\bar{w}_2\|_2^2 = 1 - \|\bar{w}\|_2^2.$$

We will assume that we are working with deflated systems: $\bar{w}_j \neq 0$ for $1 \leqq j \leqq n$, $\alpha \neq 0$, $\bar{\tau} \neq 0$, and all singular values are distinct; if not, apply the eigensystem deflation procedure [1]. Assume, then, that $\tilde{\Sigma} = \operatorname{diag}(\tilde{\sigma}_1, \ldots, \tilde{\sigma}_{n+1})$ has been found.

4.2. Calculating $\tilde{V}$

To find $\tilde{V}$, consider

$$\tilde{A}^T\tilde{A}=\begin{bmatrix} A^TA & A^Ta \\ a^TA & a^Ta \end{bmatrix}.$$

Perform an orthogonal similarity transformation by $Q=\begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix}$ to obtain

$$Q^T\tilde{A}^T\tilde{A}Q=\begin{bmatrix} \Sigma^2 & \Sigma U^Ta \\ (\Sigma U^Ta)^T & a^Ta \end{bmatrix}=\begin{bmatrix} \Sigma^2 & (1/\alpha)\Sigma\bar{w} \\ (1/\alpha)\bar{w}^T\Sigma & 1/\alpha^2 \end{bmatrix},$$

where $\alpha=1/\|a\|_2$ and $\bar{w}=\alpha U^Ta$. Then

$$\tilde{v}_i=\eta_i\begin{bmatrix} (1/\alpha)VT_i^{-1}\Sigma\bar{w} \\ -1 \end{bmatrix}, \quad \text{where} \quad \eta_i=1\Bigg/\left\|\begin{bmatrix} (1/\alpha)T_i^{-1}\Sigma\bar{w} \\ -1 \end{bmatrix}\right\|_2.$$

4.3. Calculating $\tilde{U}$

If $\tilde{A}$ is available, then $\tilde{u}_i=(1/\tilde{\sigma}_i)\tilde{A}\tilde{v}_i$ for $1\leq i\leq n$. If $\tilde{A}$ is not available, then we can compute $\tilde{u}_i$ from

$$\begin{aligned}\tilde{u}_i&=(1/\tilde{\sigma}_i)\tilde{A}\tilde{v}_i=(\eta_i/\tilde{\sigma}_i)[U\Sigma V^T \vdots a]\begin{bmatrix} (1/\alpha)VT_i^{-1}\Sigma\bar{w} \\ -1 \end{bmatrix}\\ &=(\eta_i/\tilde{\sigma}_i)[(1/\alpha)U\Sigma T_i^{-1}\Sigma\bar{w}-a]\\ &=(\eta_i/\tilde{\sigma}_i)[(1/\alpha)U\Sigma^2T_i^{-1}\bar{w}-a].\end{aligned}$$

5. Updating the SVD when Deleting a Column

Suppose we wish to delete the k^{th} column of the matrix $A\in\mathbb{R}^{m\times n}$ whose singular value decomposition is known. As usual, we let

$$\begin{aligned}A&=U\Sigma V^T, \quad U^TU=I, \quad VV^T=V^TV=I,\\ &=\hat{U}\hat{\Sigma}V^T, \quad \hat{U}^T\hat{U}=\hat{U}\hat{U}^T=I,\end{aligned}$$

$$\hat{\Sigma}=\begin{bmatrix} \Sigma \\ 0 \end{bmatrix}, \quad \hat{U}=(U,U_2),$$

$A\in\mathbb{R}^{m\times n}$, $U\in\mathbb{R}^{m\times n}$, $U_2\in\mathbb{R}^{m\times(m-n)}$, $V\in\mathbb{R}^{n\times n}$, $\Sigma\in\mathbb{R}^{n\times n}$, and $\hat{\Sigma}\in\mathbb{R}^{m\times n}$.

We may assume without loss of generality that we are deleting the last column, since permuting the columns of A is equivalent to permuting the rows of V. Therefore, let $A=(\tilde{A},a)$. We wish to calculate the SVD of $\tilde{A}$ given the SVD of A.

5.1. *Calculating $\tilde{\Sigma}$ and $\tilde{U}$*

Consider $AA^T = (\tilde{A}, a)\begin{pmatrix}\tilde{A}^T \\ a^T\end{pmatrix} = \tilde{A}\tilde{A}^T + aa^T$. Then $\tilde{A}\tilde{A}^T = AA^T - aa^T$. Premultiplying by $\hat{U}^T$ and postmultiplying by $\hat{U}$, we have $\hat{U}^T\tilde{A}\tilde{A}^T\hat{U} = \hat{\Sigma}\hat{\Sigma}^T - \hat{w}\hat{w}^T/\alpha^2$, where $\hat{w} = \alpha\hat{U}^T a$ and $\alpha = 1/\|a\|_2$. Since

$$A = (U, U_2)\begin{bmatrix}\Sigma \\ 0\end{bmatrix} V^T,$$

we have

$$\begin{bmatrix}U^T A \\ U_2^T A\end{bmatrix} = \begin{bmatrix}\Sigma V^T \\ 0\end{bmatrix},$$

which implies that $U_2^T a = 0$. Thus

$$\hat{w} = \alpha\begin{bmatrix}U^T A \\ U_2^T a\end{bmatrix} = \alpha\begin{bmatrix}U^T a \\ 0\end{bmatrix} = \begin{bmatrix}w \\ 0\end{bmatrix}.$$

As usual, we will work with deflated systems. We therefore assume that $w_j \neq 0$ for $1 \leqq j \leqq n$, and that multiple singular values are deflated out. Furthermore, a zero singular value is automatically deflated out, as the following theorem shows.

Theorem 3. Let $\tilde{A}\tilde{A}^T = AA^T - aa^T$, so that $\hat{U}^T\tilde{A}\tilde{A}^T\hat{U} = \begin{bmatrix}\Sigma^2 - ww^T/\alpha^2 & 0 \\ 0 & 0\end{bmatrix}$.

If $\sigma_k = 0$, then $w_k = 0$.

Proof. We have $A = U\Sigma V^T$ so that $U^T A = U^T(\tilde{A}, a) = (U^T\tilde{A}, U^T a) = \Sigma V^T$. If $\sigma_k = 0$, then the k^{th} row of ΣV^T is zero since Σ scales the rows of V^T. Thus, $\frac{1}{\alpha} w_k$, which is the k^{th} component of $U^T a$, is zero. QED.

The updated singular values are the roots of the secular equation:

$$f(\lambda) = 1 - (1/\alpha^2)\sum_{j=1}^{n} \frac{w_j^2}{(\sigma_j^2 - \lambda)}.$$

As in the eigensystem update [1], the new left singular vectors are given by

$$\tilde{u}_i = \gamma_i U T_i^{-1} w, \qquad \text{for } 1 \leqq i \leqq n-1,$$

where

$$\gamma_i = 1/\|T_i^{-1} w\|_2, \qquad w = \alpha U^T a, \qquad \alpha = 1/\|a\|_2, \quad \text{and} \quad T_i = \Sigma^2 - \tilde{\sigma}_i^2 I.$$

The cost to obtain $\tilde{U} \in \mathbb{R}^{m \times (n-1)}$ is $mn(n-1) + O(n^2)$ multiplications.

5.2. *Calculating $\tilde{V}$*

Since $\tilde{A} = \tilde{U}\tilde{\Sigma}\tilde{V}^T$ and $A = [\tilde{A}\,|\,a] = U\Sigma V^T$, we have

$$\tilde{A} = U\Sigma V^T\begin{bmatrix}I_{n-1} \\ \hline 0\end{bmatrix},$$

where I_{n-1} is the identity matrix of order $n-1$, and

$$A e_n = a = U \Sigma V^T e_n,$$

where e_k is the k^{th} column of I_n, $1 \leqq k \leqq n$. Thus,

$$U^T a = \Sigma V^T e_n$$

and

$$\tilde{\Sigma} \tilde{V}^T = \tilde{U}^T \tilde{A} = \tilde{U}^T U \Sigma V^T \left[\frac{I_{n-1}}{0}\right],$$

yielding

$$\tilde{v}_i = (1/\tilde{\sigma}_i) [I_{n-1} | 0]\, V \Sigma U^T \tilde{U} e_i = (1/\tilde{\sigma}_i) [I_{n-1} | 0]\, V \Sigma U^T \tilde{u}_i$$

for $1 \leqq i \leqq n$. But, from Section 5.1,

$$\tilde{u}_i = \gamma_i U T_i^{-1} w,$$

where $\gamma_i = 1/\|T_i^{-1} w\|_2$, $w = \alpha U^T a$, $\alpha = 1/\|a\|_2$, and $T_i = \Sigma^2 - \tilde{\sigma}_i I$. Hence,

$$\begin{aligned} \tilde{v}_i &= (\gamma_i/\tilde{\sigma}_i)\, V_{n-1} \Sigma T_i^{-1} w = (\alpha \gamma_i/\tilde{\sigma}_i)\, V_{n-1} \Sigma T_i^{-1} U^T a \\ &= (\alpha \gamma_i/\tilde{\sigma}_i)\, V_{n-1} \Sigma T_i^{-1} \Sigma V^T e_n, \end{aligned}$$

where V_{n-1} are the first $n-1$ rows of V. Thus, we can compute $\tilde{V} \in \mathbb{R}^{(n-1) \times (n-1)}$ with $n^3 + 2n^2 + O(n)$ multiplications.

6. Updating Least Squares Solutions

Obtaining least squares solutions to a system of overdetermined linear equations is a problem of great importance [5]. However, once having obtained a least squares solution to one system, one often seeks a least squares solution to the system obtained by appending or deleting an equation or an unknown.

6.1. Appending an Equation

Let us consider the problem of appending a row to A and an element to the right-hand side b in the least squares problem $Ax - b$. Suppose we know V, Σ and c $(= U^T b)$ and we wish to calculate $\tilde{V}$, $\tilde{\Sigma}$ and $\tilde{c}$ $(= \tilde{U}^T \tilde{b})$ for the least squares problem $\begin{bmatrix} A \\ a^T \end{bmatrix} x - \begin{bmatrix} b \\ b_{m+1} \end{bmatrix} \equiv \tilde{A} x - \tilde{b}$.

We have seen how to calculate $\tilde{V}$ and $\tilde{\Sigma}$ from V and Σ. From Section 2.2 the updated left singular vectors $\tilde{u}_i$ are given by $\tilde{u}_i = (1/\tilde{\sigma}_i) \tilde{A} \tilde{v}_i$ if $\tilde{A}$ and $\tilde{V}$ are available.

If $\tilde{A}$ or $\tilde{V}$ is not available, then

$$\tilde{u}_i = (\beta_i/\tilde{\sigma}_i) \begin{bmatrix} U T_i^{-1} \Sigma z \\ -\alpha \end{bmatrix}, \quad \text{where} \quad \beta_i = 1/\| T_i^{-1} z \|_2, \qquad z = \alpha V^T a,$$

$\alpha = 1/\|a\|_2$, and $T_i = \Sigma^2 - \tilde{\sigma}_i^2 I$. Note that we do not need U in order to calculate the β_i.

The updated right-hand side $\tilde{c} = (\tilde{c}_1, \ldots, \tilde{c}_n)^T$ is defined as $\tilde{c} \equiv \tilde{U}^T \tilde{b} = \tilde{U}^T \begin{bmatrix} b \\ b_{m+1} \end{bmatrix}$, so

$$\begin{aligned} \tilde{c}_i &= \tilde{u}_i^T \begin{bmatrix} b \\ b_{m+1} \end{bmatrix} = (\beta_i/\tilde{\sigma}_i)(z^T \Sigma T_i^{-1} U^T b - \alpha b_{m+1}) \\ &= (\beta_i/\tilde{\sigma}_i)(z^T \Sigma T_i^{-1} c - \alpha b_{m+1}). \end{aligned}$$

Thus we can compute the components of $\tilde{c}$ given only V, Σ, $\tilde{\Sigma}$ and the old right-hand side c. The calculation of each component requires $3n$ multiplications, so we obtain $\tilde{c}$ in $3n^2 + O(n)$ multiplications.

6.2. *Deleting an Equation*

Let us consider the problem of deleting the last row of A and the last element of the right-hand side b in the least squares problem $Ax - b$. Suppose we know V, Σ, and $c = U^T b$, and we wish to calculate $\tilde{V}$, $\tilde{\Sigma}$, and $\tilde{c} = \tilde{U}^T \tilde{b}$ for the least squares problem $\tilde{A} y - \tilde{b}$, where $b = (\tilde{b}, b_m)^T$.

We have seen how to calculate $\tilde{V}$ and $\tilde{\Sigma}$ from V and Σ. The updated left singular vectors $\tilde{u}_i$ are given by $\tilde{u}_i = (1/\tilde{\sigma}_i) \tilde{A} \tilde{v}_i$ if $\tilde{A}$ is still available.

If $\tilde{A}$ or $\tilde{V}$ is not available, then

$$\tilde{u}_i = (\beta_i/\tilde{\sigma}_i) U_{m-1} \Sigma T_i^{-1} z,$$

where $\beta_i = 1/\| T_i^{-1} z \|_2$, $z = \alpha V^T a$, $\alpha = 1/\|a\|_2$, $T_i = \Sigma^2 - \tilde{\sigma}_i I$, and U_{m-1} are the first $m-1$ rows of U. The components $\tilde{c}_i$ of $\tilde{c}$ are given by

$$\begin{aligned} \tilde{c}_i &= \tilde{u}_i^T b = (\beta_i/\tilde{\sigma}_i) z^T T_i^{-1} \Sigma U_{m-1}^T \tilde{b} \\ &= (\beta_i/\tilde{\sigma}_i) z^T T_i^{-1} \Sigma (c - b_m u), \end{aligned}$$

where u^T is the last row of U.

Thus we can calculate the updated right-hand side given only Σ, V, c, and $\tilde{\Sigma}$. The cost to update least squares solutions is $3n^2 + O(n)$ multiplications. The cost to calculate $\tilde{\Sigma}$ and $\tilde{V}$ is $n^3 + O(n^2)$ multiplications.

6.3. *Appending an Unknown*

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

$$\tilde{A}^T \tilde{A} = \begin{bmatrix} A^T A & A^T a \\ (A^T a)^T & a^T a \end{bmatrix} \equiv \begin{bmatrix} V \Sigma^2 V^T & r \\ r^T & 1/\alpha^2 \end{bmatrix}.$$

If we assume that we are not storing U, but that we have a mechanism for calculating $r = A^T a$, and if b is available, then we can update the least squares problem $Ax - b$.

Perform an orthogonal similarity transformation by $Q = \begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix}$:

$$Q^T \tilde{A}^T \tilde{A} Q = \begin{bmatrix} \Sigma^2 & s \\ s^T & 1/\alpha^2 \end{bmatrix}, \qquad s = V^T r = V^T A^T a.$$

We then solve the following secular equation

$$g(\lambda) = 1/\alpha^2 - \lambda - \sum_{j=1}^{n} \frac{\zeta_j^2}{(\sigma_j^2 - \lambda)} \qquad \text{for } \tilde{\sigma}_1^2, \ldots, \tilde{\sigma}_{n+1}^2.$$

We, of course, assume that we are working with deflated systems: $\zeta_j^2 \neq 0$ for $1 \leqq j \leqq n$, $\alpha \neq 0$, σ_j are distinct. Then the $n+1$ right singular vectors are

$$\tilde{v}_i = \eta_i \begin{bmatrix} V T_i^{-1} z \\ -1 \end{bmatrix}, \quad \text{where} \quad \eta_i = 1 \Big/ \left\| \begin{bmatrix} T_i^{-1} z \\ -1 \end{bmatrix} \right\|_2.$$

The updated right-hand side is $\tilde{c} = \tilde{U}^T b$, and the old right-hand side is $c = U^T b$. The components of $\tilde{c}$, for $1 \leqq i \leqq n+1$, are given by

$$\begin{aligned} \tilde{c}_i &= \tilde{u}_i^T b \\ &= (\eta_i/\tilde{\sigma}_i)[(1/\alpha) U \Sigma^2 T_i^{-1} \bar{w} - a]^T b \\ &= (\eta_i/\tilde{\sigma}_i)[(1/\alpha) \bar{w}^T T_i^{-1} \Sigma^2 U^T b - a^T b] \\ &= (\eta_i/\tilde{\sigma}_i)[(1/\alpha) a^T U \Sigma^2 T_i^{-1} c - a^T b]. \end{aligned}$$

Since $V^T A^T a = \Sigma U^T a = s$, we have $\tilde{c}_i = (\eta_i/\tilde{\sigma}_i)[(1/\alpha) s^T \Sigma T_i^{-1} c - a^T b]$.

Thus, least squares solutions can be updated, although not as efficiently as when adding or deleting rows since $A^T a$ and b are required; but that is the price one pays for adding a column part way through the updating process.

6.4. Deleting an Unknown

Let $A = (\tilde{A}, a)$. Then $\tilde{A}\tilde{A}^T = AA^T - aa^T$.

If U is not being stored, then we assume that there is a mechanism for calculating $A^T a$.

But

$$\hat{U}^T \tilde{A} \tilde{A}^T \hat{U} = \hat{\Sigma} \hat{\Sigma}^T - \frac{1}{\alpha^2} \hat{w} \hat{w}^T = \begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix} - (1/\alpha^2) \begin{bmatrix} w w^T & 0 \\ 0 & 0 \end{bmatrix},$$

where

$$\hat{w} = \alpha \hat{U}^T a = \begin{bmatrix} U^T a \\ 0 \end{bmatrix} = \begin{bmatrix} w \\ 0 \end{bmatrix}, \quad \text{and} \quad \alpha = 1/\|a\|_2.$$

Thus, if $w = \alpha U^T a$, then $V \Sigma w = \alpha V \Sigma U^T a = \alpha A^T a$, and $w = \alpha \Sigma^{-1} V^T A^T a$.

Since we are working with deflated systems, we are assuming $\sigma_k > 0$ for $1 \leqq k \leqq n$. One should rescale the quantity $\Sigma^{-1} V^T A^T a$ so that its 2-norm is 1, and compute $\alpha = 1/\|a\|_2$.

Once w is computed, we may find the updated singular values by solving the secular equation

$$f(\lambda) = 1 - (1/\alpha^2) \sum_{j=1}^{n} \frac{\zeta_j^2}{(\sigma_j^2 - \lambda)}.$$

The updated right singular vectors $\tilde{v}_i$ are given as in Section 5.2:

$$\tilde{v}_i = (\alpha \gamma_i / \tilde{\sigma}_i) V_{n-1} \Sigma T_i^{-1} \Sigma V^T e_n \qquad \text{for } 1 \leqq i \leqq n-1$$

where $\gamma_i = 1/\|T_i^{-1} w\|_2$, $w = U^T a$, $\alpha = 1/\|a\|_2$, $T_i = \Sigma^2 - \tilde{\sigma}_i^2 I$, and V_{n-1} is the first $n-1$ rows of V.

The updated right-hand side $\tilde{c}$ is given by $\tilde{c} = \tilde{U}^T b$. If $\tilde{c} = (\tilde{c}_1, \ldots, \tilde{c}_{n-1})$, we have $\tilde{c}_i = \tilde{u}_i^T b$ for $1 \leqq i \leqq n-1$.

From Section 5.1, $\tilde{u}_i = \gamma_i U T_i^{-1} w$. Thus,

$$\tilde{c}_i = \gamma_i w^T T_i^{-1} U^T b = \gamma_i w^T T_i^{-1} c,$$

where $c = U^T b$ is the old right-hand side.

The cost for updating least squares solutions is $2n^2 + O(n)$ multiplications. The cost for calculating $\tilde{V}$ and $\tilde{\Sigma}$ is $n^3 + O(n^2)$ multiplications.

7. Numerical Results

For problem (I a), three examples were run using a FORTRAN code on a PDP-11/40. The machine precision is $\eta = 2^{-23}$. The rows used in updating came from the rows of the $m \times n$ Hilbert section $H(m,n)$, whose (i,j) element is given by $1/(i+j-1)$. The i^{th} row is denoted by h_i^T.

For the first example we started with a 5×5 system, where $\Sigma = \text{diag}(1, 2, 2, 2, 2)$, and $U = V = I$. We then appended 15 rows of $20 \times H(15,5)$, one at a time, i.e., $A^{(0)} = \Sigma$,

$$A^{(k)} = \begin{bmatrix} A^{(k-1)} \\ 20 h_k^T \end{bmatrix}.$$

The results are given in Table 1, Ex. 1.

The second example began with $\Sigma = \text{diag}(0,0,0,0,0)$, $U = V = I$. We then appended 15 rows from $H(15,5)$. The results are given in Table 1, Ex. 2.

The last example began with $\Sigma = 0$, $U = V = I$, all 10×10. We then appended 30 rows from $H(30,10)$. The results are given in Table 1, Ex. 3.

A few remarks are in order. From Ex. 1 through Ex. 3, the problem becomes progressively more ill-conditioned (with respect to solving least squares problems). The 2-condition numbers for the examples are 17, 52500, and 2×10^8, respectively. We consider the results given in Table 1 very good, especially in Example 3. Note that in every example the code was required to perform

Table 1. Norms are in units of the machine precision, $\eta = 2^{-23}$

	m	$\|I - V^T V\|_1$	$\|I - U^T U\|_1$	$\frac{\|A - U\Sigma V^T\|_1}{\|A\|_1}$
Ex. 1:	6	4	3	0.2
	10	5	3	1.3
	15	10	5	1.3
	20	12	10	1.9
Ex. 2:	6	1	1	1.0
	10	9	4	2.0
	15	14	5	2.0
	20	18	10	2.0
Ex. 3:	11	1	1	0.5
	15	10	5	1.25
	20	15	10	1.7
	25	24	16	2.4
	30	34	24	4.0
	35	45	26	1.3
	40	56	35	1.3

deflation. We set the deflation tolerance to zero so that a group of singular values were deflated only if they were (numerically) equal. We feel that this is a good indication that the algorithm for computing the updated singular values is robust, even in the presence of quite close singular values.

For the least squares problem (II), we tested the code by using right-hand sides $e_1, e_2, \dots, e_m$. This should produce, for each e_i, the i^{th} row of U. Our tests show that this is, in fact, the case to within 2 machine units. Since the method for computing $\tilde{V}$ and $\tilde{\Sigma}$ remains the same, the results in Table 1 also apply here.

8. Remarks

The case of a rank-one modification to the singular value decomposition, $\tilde{A} = A + uv^T$, is not treated here. The updating problems that we have considered were transformed into rank-one modifications of a symmetric eigenvalue problem. A rank-one modification to the SVD is actually a rank-two modification of the associated eigenproblem (cf. Thompson [8]). Similarly, Thompson [8] shows that a rank-k modification to the SVD is a rank-$2k$ modification to the associated eigenproblem. The method in [1] for computing a rank-one update to the symmetric eigenproblem does not readily extend to the rank-k case, although a rank-k update can always be computed by a sequence of rank-one updates. For this reason, our algorithm for updating the SVD does not extend in an obvious way to the rank-one update of the SVD. However, the rank-two symmetric eigenproblem associated with the rank-one SVD update is "special"; it may be possible to use the additional structure to devise an algorithm to treat this special case.

Acknowledgments. We would like to thank Linda Kaufman for her excellent suggestions, which helped to simplify the derivation and presentation. We would also like to thank the referees for their helpful comments.

References

1. Bunch, J.R., Nielsen, C.P., Sorensen, D.C.: Rank-one modification of the symmetric eigenproblem. Numer. Math. **31**, 31-48 (1978)
2. Businger, P.A.: Updating a singular value decomposition. Nordisk Tidskr. Informations behandling (BIT) **10**, 376-397 (1970)
3. Golub, G.H.: Some modified matrix eigenvalue problems. SIAM Rev. **15**, 318–334 (1973)
4. Jennings, L.S., Osborne, M.R.: A direct error analysis for least squares. Numer. Math. **22**, 325–332 (1974)
5. Lawson, C.L., Hanson, R.J.: Solving least squares problems. New Jersey: Prentice-Hall 1974
6. Osborne, M.R.: On the computation of stepwise regressions. Austral. Comput. J. **8**, 61–68 (1976)
7. Stewart, G.W.: Introduction to matrix computations. New York: Academic Press 1973
8. Thompson, R.C.: The behavior of eigenvalues and singular values under perturbations of restricted rank. Linear Algebra and Appl. **13**, 69–78 (1976)
9. Wilkinson, J.H.: The algebraic eigenvalue problem. Oxford: Clarendon Press 1965

Received June 7, 1976