Aplikace matematiky

Gene Howard Golub

# Least squares, singular values and matrix approximations

*Aplikace matematiky*, Vol. 13 (1968), No. 1, 44–51

Persistent URL: http://dml.cz/dmlcz/103138

**Terms of use:**

© Institute of Mathematics AS CR, 1968

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://dml.cz

# LEAST SQUARES, SINGULAR VALUES AND MATRIX APPROXIMATIONS[1])

GENE HOWARD GOLUB

**0.** Let $A$ be a real, $m \times n$ matrix (for notational convenience we assume that $m \geqq n$). It is well known (cf. [6]) that

$$A = U\Sigma V^T \tag{0.1}$$

where

$$UU^T = I_m, \quad VV^T = I_n$$

and

$$\Sigma = \begin{pmatrix} \sigma_1, \ldots, 0 \\ \ldots\ldots\ldots \\ 0, \quad \ldots, \sigma_n \\ \hline 0 \end{pmatrix} \} \; (m - n) \times n \, .$$

The matrix $U$ consists of the orthonormalized eigenvectors of $AA^T$, and the matrix $V$ consists of the orthonormalized eigenvectors of $A^TA$. The diagonal elements of $\Sigma$ are the non-negative square roots of the eigenvalues of $A^TA$; they are called *singular values* or *principal values* of $A$. Throughout this note, we assume

$$\sigma_1 \geqq \sigma_2 \geqq \ldots \geqq \sigma_n \geqq 0 \, .$$

Thus if $\operatorname{rank}(A) = r$, $\sigma_{r+1} = \sigma_{r+2} = \ldots = \sigma_n = 0$. The decomposition (0.1) is called the *singular value decomposition*.

In this paper, we shall present a numerical algorithm for computing the singular value decomposition.

**1.** The singular value decomposition plays an important role in a number of least squares problems, and we will illustrate this with some examples. Throughout this discussion, we use the euclidean or Frobenius norm of a matrix, viz.

$$\|A\| = \left(\sum_{i,j} |a_{ij}|^2\right)^{1/2} \, .$$

[1]) This work was in part supported by NSF and ONR.

A. Let $\mathcal{U}_n$ be the set of all $n \times n$ orthogonal matrices. For an arbitrary $n \times n$ real matrix $A$, determine $Q \in \mathcal{U}_n$ such that

$$\|A - Q\| \leqq \|A - X\| \quad \text{for any} \quad X \in \mathcal{U}_n .$$

It has been shown by FAN and HOFFMAN [2] that if

$$A = U\Sigma V^T , \quad \text{then} \quad Q = UV^T .$$

B. An important generalization of problem A occurs in factor analysis. For arbitrary $n \times n$ real matrices $A$ and $B$, determine $Q \in \mathcal{U}_n$ such that

$$\|A - BQ\| \leqq \|A - BX\| \quad \text{for any} \quad X \in \mathcal{U}_n .$$

It has been shown by GREEN [5] and by SCHÖNEMANN [9] that if

$$B^T A = U\Sigma V^T , \quad \text{then} \quad Q = UV^T .$$

C. Let $\mathcal{M}_{m,n}^{(k)}$ be the set of all $m \times n$ matrices of rank $k$. Assume $A \in \mathcal{M}_{m,n}^{(r)}$. Determine $B \in \mathcal{M}_{m,n}^{(k)}$ $(k \leqq r)$ such that

$$\|A - B\| \leqq \|A - X\| \quad \text{for all} \quad X \in \mathcal{M}_{m,n}^{(k)} .$$

It has been shown by ECKART and YOUNG [1] that if

$$(1.1) \qquad A = U\Sigma V^T , \quad \text{then} \quad B = U\Omega_k V^T$$

where

$$(1.2) \qquad \Omega_k = \left( \begin{array}{c|c} \begin{matrix} \sigma_1, & 0, & \ldots, & 0 \\ 0, & \sigma_2, & \ldots, & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0, & 0, & \ldots, & \sigma_k \end{matrix} & \\ \hline & 0 \end{array} \right) .$$

Note that

$$(1.3) \qquad \|A - B\| = \|\Sigma - \Omega_k\| = (\sigma_{k+1}^2 + \ldots + \sigma_r^2)^{1/2} .$$

D. An $n \times m$ matrix $X$ is said to be the *pseudo-inverse* of an $m \times n$ matrix $A$ if $X$ satisfies the following four properties:

i) $AXA = A$,
ii) $XAX = X$,
iii) $(AX)^T = AX$,
iv) $(XA)^T = XA$.

We denote the pseudo-inverse by $A^+$. We wish to determine $A^+$ numerically. It can be shown [8] that $A^+$ can always be determined and is unique.

It is easy to verify that

$$A^{+} = V\Lambda U^{T} \tag{1.4}$$

where

$$\Lambda = \left[\begin{array}{cccc|c} \frac{1}{\sigma_1}, & 0, & \ldots, & 0 & \\ 0, & \frac{1}{\sigma_2}, & \ldots, & 0 & \\ \multicolumn{4}{c|}{\cdots\cdots\cdots\cdots} & \\ 0, & 0, & \ldots, & \frac{1}{\sigma_r} & \\ \hline & & & & 0 \end{array}\right] n \times m\,.$$

In recent years there have been a number of algorithms proposed for computing the pseudo-inverse of a matrix. These algorithms usually depend upon a knowledge of the rank of the matrix or upon some suitably chosen parameter. For example in the latter case, if one uses (1.4) to compute the pseudo-inverse, then after one has computed the singular value decomposition numerically it is necessary to determine which of the singular values are zero by testing against some tolerance.

Alternatively, suppose we know that the given matrix $A$ can be represented as

$$A = B + \delta B$$

where $\delta B$ is a matrix of perturbations and

$$\|\delta B\| \leqq \eta\,.$$

Now, we wish to construct a matrix $\hat{B}$ such that

$$\|A - \hat{B}\| \leqq \eta$$

and

$$\operatorname{rank}(\hat{B}) = \text{minimum}\,.$$

This can be accomplished with the aid of the solution to problem (C). Let

$$B_k = U\Omega_k V^T$$

as in equation (1.2).

Then using (1.3),

$$\hat{B} = B_p$$

if

$$(\sigma_{p+1}^2 + \sigma_{p+2}^2 + \ldots + \sigma_n^2)^{1/2} \leqq \eta$$

and

$$(\sigma_p^2 + \sigma_{p+1}^2 + \ldots + \sigma_n^2)^{1/2} > \eta\,.$$

Since rank $(\hat{B}) = p$ by construction,

$$\hat{B}^+ = V\Omega_p^+ U^T .$$

Thus, we take $\hat{B}^+$ as our approximation to $A^+$.

E. Let $A$ be a given matrix, and let $b$ be a known vector. Determine a vector $x$ such that for

$$\|b - Ax\|_2 = \min \quad \text{and} \quad \|x\|_2 = \min ,$$

where $\|y\|_2 = (\sum y_i^2)^{1/2}$ for any vector $y$. It is easy to verify that $x = A^+ b$.

A norm is said to be *unitarily invariant* if $\|AU\| = \|VA\| = \|A\|$ when $U^*U = I$ and $V^*V = I$. FAN and HOFFMAN [2] have shown that the solution to problem (A) is the same for all unitarily invariant norms and MIRSKY [7] has proved a similar result for the solution to problem (C).

**2.** In [4] it was shown by GOLUB and KAHAN that it is possible to construct a sequence of orthogonal matrices $\{P^{(k)}\}_{k=1}^n$, $\{Q^{(k)}\}_{k=1}^{n-1}$ via Householder transformation so that

$$P^{(n)}P^{(n-1)} \ldots P^{(1)}AQ^{(1)}Q^{(2)} \ldots Q^{(n-1)} \equiv P^TAQ = J$$

and $J$ is an $m \times n$ bi-diagonal matrix of the form

$$J = \left[\begin{array}{ccccc} \alpha_1, & \beta_1, & 0, & \ldots, & 0 \\ 0, & \alpha_2, & \beta_2 & \ldots, & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0, & 0, & 0, & \ldots, & \beta_{n-1} \\ 0, & 0, & 0, & \ldots, & \alpha_n \\ \hline & & 0 & & \end{array}\right] \} (m-n) \times n .$$

The singular values of $J$ are the same as those of $A$. Thus if the singular value decomposition of

$$J = X\Sigma Y^T$$

then

$$A = PX\Sigma Y^T Q^T$$

so that $U = PX$, $V = QY$.

A number of algorithms were proposed in [4] for computing the singular value decomposition of $J$. We now describe a new algorithm, based on the $QR$ algorithm of FRANCIS [3], for computing the singular value decomposition of $J$.

Let

$$K = \begin{bmatrix} 0 & \alpha_1 & & & & & \\ \alpha_1 & 0 & \beta_1 & & & 0 & \\ & \beta_1 & 0 & \alpha_2 & & & \\ & & \alpha_2 & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & 0 & & & \cdot & \cdot & \alpha_n \\ & & & & & \alpha_n & 0 \end{bmatrix} 2n \times 2n .$$

It can be shown [4] that $K$ is a symmetric, tri-diagonal matrix whose eigenvalues are $\pm$ singular values of $J$. One of the most effective methods of computing the eigenvalues of a tri-diagonal matrix is the $QR$ algorithm of Francis, which proceeds as follows:

Begin with the given matrix $K = K_0$. Compute the factorization

$$K_0 = M_0 R_0$$

where $M_0^T M_0 = I$ and $R_0$ is an upper triangular matrix, and then multiply the matrices in reverse order so that

$$K_1 = R_0 M_0 = M_0^T K_0 M_0 .$$

Now one treats $K_1$ in the same fashion as the matrix $K_0$, and a sequence of matrices is obtained by continuing *ad infinitum.* Thus

$$K_i = M_i R_i \quad \text{and} \quad K_{i+1} = R_i M_i = M_{i+1} R_{i+1} ,$$

so that

$$K_{i+1} = M_i^T K_i M_i = M_i^T M_{i-1}^T \dots M_0^T K M_0 M_1 \dots M_i .$$

The method has the advantage that $K_i$ remains tri-diagonal throughout the computation.

For suitably chosen *shift parameters* $s_i$, we can accelerate the convergence of the $QR$ method by computing

$$(2.1) \qquad (K_i - s_i I) = M_i R_i , \quad R_i M_i + s_i I = K_{i+1} .$$

Unfortunately, the shift parameter $s_i$ may destroy the zeroes on the diagonal of $K$.

Since the eigenvalues of $K$ always occur in pairs, it would seem more appropriate to compute the $QR$ decomposition of

$$(K_i - s_i I)(K_i + s_i I) = K_i^2 - s_i^2 I$$

so that

$$M_i R_i = K_i^2 - s_i^2 I .$$

It has been shown by Francis that it is not necessary to compute (2.1) explicitly but it is possible to perform the shift implicitly. Let

$$\{N_i\}_{k,1} = \{M_i\}_{k,1} , \quad k = 1, 2, \dots, 2n .$$

(i.e., the elements of the first column of $N$ are equal to the elements of the first column of $M$) and

$$N_i^T N_i = I .$$

Then if

i) $T_{i+1} = N_i^T K_i N_i$,
ii) $T_{i+1}$ is a tri-diagonal matrix,

iii) $K_i$ is non-singular,
iv) the sub-diagonal elements of $T_{i+1}$ are positive,

it follows that $T_{i+1} = K_{i+1}$.

The calculation proceeds quite simply. Dropping the iteration counter i), let

$$
Z_p = \begin{array}{c}
\begin{array}{ccccc} & (p) & (p+1) & (p+2) & \end{array} \\
\begin{bmatrix}
1 & & & & & & & \\
 & \ddots & & & & & 0 & \\
 & & 1 & & & & & \\
 & & & \cos\theta_p, & 0, & \sin\theta_p & & \\
 & & & 0, & 1, & 0 & & \\
 & & & \sin\theta_p, & 0, & -\cos\theta_p & & \\
 & & & & & & 1 & \\
 & 0 & & & & & & \ddots \\
 & & & & & & & & 1
\end{bmatrix}
\begin{array}{l} \\ \\ \\ (p) \\ (p+1) \\ (p+2) \\ \\ \\ 2n \times 2n\,. \end{array}
\end{array}
$$

Then $\cos\theta_1$ is chosen so that

$$\{Z_1(K^2 - s^2 I)\}_{k,1} = 0 \quad \text{for} \quad k = 2, 3, \ldots, 2n\,.$$

Then the matrix

$$
Z_1 K Z_1 = \begin{bmatrix}
0 & \alpha_1' & 0 & d_1 & & & \\
\alpha_1' & 0 & \beta_1' & & & & 0 \\
0 & \beta_1' & \cdot & \alpha_2' & & & \\
d_1 & & \alpha_2' & \cdot & \beta_2 & & \\
 & & & \beta_2 & \cdot & \cdot & \\
 & 0 & & & \cdot & \cdot & \alpha_n \\
 & & & & \cdot & \alpha_n & 0
\end{bmatrix};
$$

and

$$T = Z_{2n-2} \ldots Z_1 K Z_1 \ldots Z_{2n-2}\,,$$

where $Z_2, \ldots, Z_{2n-2}$ are constructed so that $T$ is tri-diagonal. The product of all the orthogonal transformations which gives the singular values yields the matrix of orthogonal eigenvectors of $K$. For ease of notation let us write

$$\gamma_{2j-1} = \alpha_j\,, \quad j = 1, 2, \ldots, n\,,$$
$$\gamma_{2j} = \beta_j\,, \quad j = 1, 2, \ldots, n-1\,.$$

Then explicitly, the calculation goes as follows: Dropping the iteration counter $i$,

$$\gamma_0 = \gamma_1^2 - s^2\,, \quad d_0 = \gamma_1\gamma_2\,.$$

For $j = 0, 1, \ldots, 2n-3$,

$$\begin{aligned}
r_j &= (\gamma_j^2 + d_j^2)^{1/2}, \\
\sin\theta_j &= d_j/r_j, \quad \cos\theta_j = \gamma_j/r_j, \\
\gamma_j &= r_j, \\
\bar{\gamma}_{j+1} &= \tilde{\gamma}_{j+1}\cos\theta_j + \hat{\gamma}_{j+2}\sin\theta_j, \\
\tilde{\gamma}_{j+2} &= \tilde{\gamma}_{j+1}\sin\theta_j - \hat{\gamma}_{j+2}\cos\theta_j, \\
\hat{\gamma}_{j+3} &= -\gamma_{j+3}\cos\theta_j, \\
d_{j+1} &= \phantom{-}\gamma_{j+3}\sin\theta_j.
\end{aligned}$$

In the actual computation, no additional storage is required for

$$\{\bar{\gamma}_j, \tilde{\gamma}_j, \hat{\gamma}_j\}$$

since they may overwrite $\{\gamma_j\}$. Furthermore, only one element of storage need be reserved for $\{d_j\}$. When $|\gamma_{2n-2}|$ is sufficiently small, $|\gamma_{2n-1}|$ is taken as a singular value and $n$ is replaced by $n-1$.

Now let us define

$$W_p = \begin{array}{cc}
\begin{array}{cccccc} \phantom{1} & \phantom{\ddots} & [p/2+1] & [p/2+2] & \phantom{\ddots} & \phantom{1} \end{array} & \\
\left[\begin{array}{cccccc}
1 & & & & & \\
 & \ddots & & & 0 & \\
 & & 1 & & & \\
 & & \cos\theta_p & \sin\theta_p & & \\
 & & \sin\theta_p & -\cos\theta_p & & \\
 & & & 1 & & \\
 0 & & & & \ddots & \\
 & & & & & 1
\end{array}\right] &
\begin{array}{l} \\ \\ \\ [p/2+2] \\ [p/2+1] \\ \\ \\ n \times n \end{array}
\end{array}$$

where $\cos\theta_p$ is defined as above. It has been pointed out to the author by J. H. WILKINSON that the above iteration is equivalent to forming

$$\hat{J} = W_{2n-3} \ldots W_3 W_1 J W_0 W_2 \ldots W_{2n-2}$$

where $\hat{J}$ is again a bi-diagonal matrix. Thus,

$$\begin{aligned}
X &= \prod_i (W_1^{(i)} W_3^{(i)} \ldots W_{2n-3}^{(i)}), \\
Y &= \prod_i (W_0^{(i)} W_2^{(i)} \ldots W_{2n-2}^{(i)}).
\end{aligned}$$

An ALGOL procedure embodying these techniques will soon be published by Dr. PETER BUSINGER and the author.

## *References*

An extensive list of references on singular values is given in [4].

[1] *C. Eckart* and *G. Young*, "The approximation of one matrix by another of lower rank", Psychometrika, *1* (1936), pp. 211—218.

[2] *K. Fan* and *A. Hoffman*, "Some metric inequalities in the space of matrices", Proc. Amer. Math. Soc., *6* (1955), pp. 111—116.

[3] *J. Francis*, "The *QR* transformation. A unitary analogue to the *LR* transformation", Comput. J., *4* (1961, 1962), pp. 265—271.

[4] *G. Golub* and *W. Kahan*, "Calculating the singular values and pseudoinverse of a matrix", J. SIAM Numer. Anal. Ser. B, *2* (1965), pp. 205—224.

[5] *B. Green*, "The orthogonal approximation of an oblique structure in factor analysis", Psychometrika, *17* (1952), pp. 429—440.

[6] *C. Lanczos*, Linear Differential Operators, Van Nostrand, London, 1961, Chap. 3.

[7] *L. Mirsky*, "Symmetric gauge functions and unitarily invariant norms", Quart. J. Math. Oxford (2), *11* (1960), pp. 50—59.

[8] *R. Penrose*, "A generalized inverse for matrices", Proc. Cambridge Philos. Soc., *51* (1955), pp. 406—413.

[9] *P. Schönemann*, "A generalized solution of the orthogonal procrustes problem", Psychometrika, *31* (1966), pp. 1—10.

*Gene Howard Golub*, Computer Science Dept., Stanford, Calif., 94305, U.S.A.