Matrix Computations

THIRD EDITION

Gene H. Golub
Department of Computer Science
Stanford University

Charles F. Van Loan
Department of Computer Science
Cornell University

The Johns Hopkins University Press
Baltimore and London

Printed in the United States of America on acid-free paper
9 8 7 6

First edition 1983
Second edition 1989
Third edition 1996

The Johns Hopkins University Press
2715 North Charles Street
Baltimore, Maryland 21218-4363
www.press.jhu.edu

Library of Congress Cataloging-in-Publication Data will be found at the end of this book.

A catalog record for this book is available from the British Library.

ISBN 0-8018-5413-X
ISBN 0-8018-5414-8 (pbk.)

undoff error, we recommend

lgebraic Processes, Prentice-Hall, Engle-

is," SIAM Review 12, 548–68.
'umerical Methods and Software, Prentice-

ctures on Finite Precision Computations,

olve interval analysis, the building of sta-
nating of the analysis itself:

Probabilistic Models for Propagation of

ulation of the Effects of Roundoff Errors,"

for Roundoff Analysis, II," ACM Trans.

metic: A Reasonable Portable Package,"

ent needs a thorough understanding of
acquiring knowledge in this direction is
l in

Scientist Should Know About Floating

sion Arithmetic Package," ACM Trans.

:an Multiple Precision Arithmetic Pack-

ability of Numerical Software," SIAM J.

e Arithmetic of the Digital Computer,"

IAR: A Subroutine to Dynamically De-
Math. Soft. 14, 303–311.
uts (1989). "Floating Point Arithmetic
mputing Appl. 3, 86–90.
sion Translation and Execution of FOR-
19, 288–319.

of high-quality software, even for "sim-
the design of a subroutine to compute

gram to Find the Euclidean Norm of a

her "fast" linear algebra procedures see

R.P. Brent (1970). "Error Analysis of Algorithms for Matrix Multiplication and Triangular Decomposition Using Winograd's Identity," *Numer. Math. 16*, 145–156.

W. Miller (1975). "Computational Complexity and Numerical Stability," *SIAM J. Computing 4*, 97–107.

N.J. Higham (1992). "Stability of a Method for Multiplying Complex Matrices with Three Real Matrix Multiplications," *SIAM J. Matrix Anal. Appl. 13*, 681–687.

J.W. Demmel and N.J. Higham (1992). "Stability of Block Algorithms with Fast Level-3 BLAS," *ACM Trans. Math. Soft. 18*, 274–291.

2.5 Orthogonality and the SVD

Orthogonality has a very prominent role to play in matrix computations. After establishing a few definitions we prove the extremely useful singular value decomposition (SVD). Among other things, the SVD enables us to intelligently handle the matrix rank problem. The concept of rank, though perfectly clear in the exact arithmetic context, is tricky in the presence of roundoff error and fuzzy data. With the SVD we can introduce the practical notion of numerical rank.

2.5.1 Orthogonality

A set of vectors $\{x_1, \ldots, x_p\}$ in $\mathbb{R}^m$ is *orthogonal* if $x_i^T x_j = 0$ whenever $i \neq j$ and *orthonormal* if $x_i^T x_j = \delta_{ij}$. Intuitively, orthogonal vectors are maximally independent for they point in totally different directions.

A collection of subspaces $S_1, \ldots, S_p$ in $\mathbb{R}^m$ is *mutually orthogonal* if $x^T y = 0$ whenever $x \in S_i$ and $y \in S_j$ for $i \neq j$. The *orthogonal complement* of a subspace $S \subseteq \mathbb{R}^m$ is defined by

$$S^\perp = \{y \in \mathbb{R}^m : y^T x = 0 \text{ for all } x \in S\}$$

and it is not hard to show that $\text{ran}(A)^\perp = \text{null}(A^T)$. The vectors $v_1, \ldots, v_k$ form an *orthonormal* basis for a subspace $S \subseteq \mathbb{R}^m$ if they are orthonormal and span S.

A matrix $Q \in \mathbb{R}^{m \times m}$ is said to be *orthogonal* if $Q^T Q = I$. If $Q = [\, q_1, \ldots, q_m \,]$ is orthogonal, then the q_i form an orthonormal basis for $\mathbb{R}^m$. It is always possible to extend such a basis to a full orthonormal basis $\{v_1, \ldots, v_m\}$ for $\mathbb{R}^m$:

Theorem 2.5.1 *If $V_1 \in \mathbb{R}^{n \times r}$ has orthonormal columns, then there exists $V_2 \in \mathbb{R}^{n \times (n-r)}$ such that*

$$V = [\, V_1 \; V_2 \,]$$

is orthogonal. Note that $\text{ran}(V_1)^\perp = \text{ran}(V_2)$.

Proof. This is a standard result from introductory linear algebra. It is also a corollary of the QR factorization that we present in §5.2. □

2.5.2 Norms and Orthogonal Transformations

The 2-norm is invariant under orthogonal transformation, for if $Q^TQ = I$, then $\| Qx \|_2^2 = x^TQ^TQx = x^Tx = \| x \|_2^2$. The matrix 2-norm and the Frobenius norm are also invariant with respect to orthogonal transformations. In particular, it is easy to show that for all orthogonal Q and Z of appropriate dimensions we have

$$\| QAZ \|_F = \| A \|_F \qquad (2.5.1)$$

and

$$\| QAZ \|_2 = \| A \|_2 . \qquad (2.5.2)$$

2.5.3 The Singular Value Decomposition

The theory of norms developed in the previous two sections can be used to prove the extremely useful singular value decomposition.

Theorem 2.5.2 (Singular Value Decomposition (SVD)) *If A is a real m-by-n matrix, then there exist orthogonal matrices*

$$U = [\, u_1, \ldots, u_m \,] \in \mathbb{R}^{m\times m} \quad \textit{and} \quad V = [\, v_1, \ldots, v_n \,] \in \mathbb{R}^{n\times n}$$

such that

$$U^TAV = \text{diag}(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{m\times n} \quad p = \min\{m, n\}$$

where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$.

Proof. Let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ be unit 2-norm vectors that satisfy $Ax = \sigma y$ with $\sigma = \| A \|_2$. From Theorem 2.5.1 there exist $V_2 \in \mathbb{R}^{n\times(n-1)}$ and $U_2 \in \mathbb{R}^{m\times(m-1)}$ so $V = [\, x \; V_2 \,] \in \mathbb{R}^{n\times n}$ and $U = [\, y \; U_2 \,] \in \mathbb{R}^{m\times m}$ are orthogonal. It is not hard to show that U^TAV has the following structure:

$$U^TAV = \begin{bmatrix} \sigma & w^T \\ 0 & B \end{bmatrix} \equiv A_1 .$$

Since

$$\left\| A_1 \left(\begin{bmatrix} \sigma \\ w \end{bmatrix} \right) \right\|_2^2 \geq (\sigma^2 + w^Tw)^2$$

we have $\| A_1 \|_2^2 \geq (\sigma^2 + w^Tw)$. But $\sigma^2 = \| A \|_2^2 = \| A_1 \|_2^2$, and so we must have $w = 0$. An obvious induction argument completes the proof of the theorem. $\square$

The σ_i are the *singular values* of A and the vectors u_i and v_i are the ith *left singular vector* and the ith *right singular vector* respectively. It

is easy to verify by compar
$A^TU = V\Sigma^T$ that

$$Av_i =$$
$$A^Tu_i =$$

It is convenient to have the f
ues:

$$\sigma_i(A) = \text{th}$$
$$\sigma_{max}(A) = \text{th}$$
$$\sigma_{min}(A) = \text{th}$$

The singular values of a matri
of the hyperellipsoid E define

Example 2.5.1

$$A = \begin{bmatrix} .96 & 1.72 \\ 2.28 & .96 \end{bmatrix} = U\Sigma V$$

The SVD reveals a great
SVD of A is given by Theorem

$$\sigma_1 \geq \cdots \geq \sigma$$

then

$$\text{rank}(A)$$
$$\text{null}(A)$$
$$\text{ran}(A)$$

and we have the *SVD expansion*

$$A$$

Various 2-norm and Frobeni
SVD. If $A \in \mathbb{R}^{m\times n}$, then

$$\| A \|_F^2 =$$
$$\| A \|_2 =$$
$$\min_{x\neq 0} \frac{\| Ax \|_2}{\| x \|_2} =$$

ransformations

ransformation, for if $Q^TQ =$ ...
$x \|_2^2$. The matrix 2-norm a...
respect to orthogonal transf...
hat for all orthogonal Q and ...

$\|_F$ (2.5.1)

$\|_2$. (2.5.2)

mposition

ous two sections can be used t...
ecomposition.

position (SVD)) *If A is a r...*
matrices

$= [\, v_1, \dots, v_n \,] \in \mathbb{R}^{n \times n}$

$n \quad p = \min\{m, n\}$

orm vectors that satisfy $Ax =$
there exist $V_2 \in \mathbb{R}^{n \times (n-1)}$ and
nd $U = [\, y \;\; U_2 \,] \in \mathbb{R}^{m \times m}$ are
V has the following structure:

$= A_1.$

$r^2 + w^T w)^2$

$\| A \|_2^2 = \| A_1 \|_2^2$, and so we
gument completes the proof of

. the vectors u_i and v_i are the
ingular vector respectively. It

is easy to verify by comparing columns in the equations $AV = U\Sigma$ and $A^TU = V\Sigma^T$ that

$$\left.\begin{array}{rcl} Av_i & = & \sigma_i u_i \\ A^T u_i & = & \sigma_i v_i \end{array}\right\} \; i = 1{:}\min\{m, n\}$$

It is convenient to have the following notation for designating singular values:

$$\begin{array}{rcl} \sigma_i(A) & = & \text{the } i\text{th largest singular value of } A, \\ \sigma_{max}(A) & = & \text{the largest singular value of } A, \\ \sigma_{min}(A) & = & \text{the smallest singular value of } A. \end{array}$$

The singular values of a matrix A are precisely the lengths of the semi-axes of the hyperellipsoid E defined by $E = \{ \, Ax : \| x \|_2 = 1 \, \}$.

Example 2.5.1

$$A = \begin{bmatrix} .96 & 1.72 \\ 2.28 & .96 \end{bmatrix} = U\Sigma V^T = \begin{bmatrix} .6 & -.8 \\ .8 & .6 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} .8 & .6 \\ .6 & -.8 \end{bmatrix}^T .$$

The SVD reveals a great deal about the structure of a matrix. If the SVD of A is given by Theorem 2.5.2, and we define r by

$$\sigma_1 \ge \cdots \ge \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0,$$

then

$$\begin{array}{rcll} \text{rank}(A) & = & r & (2.5.3) \\ \text{null}(A) & = & \text{span}\{v_{r+1}, \dots, v_n\} & (2.5.4) \\ \text{ran}(A) & = & \text{span}\{u_1, \dots, u_r\}\,, & (2.5.5) \end{array}$$

and we have the *SVD expansion*

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T . \qquad (2.5.6)$$

Various 2-norm and Frobenius norm properties have connections to the SVD. If $A \in \mathbb{R}^{m \times n}$, then

$$\begin{array}{rcll} \| A \|_F^2 & = & \sigma_1^2 + \cdots + \sigma_p^2 \qquad p = \min\{m, n\} & (2.5.7) \\ \| A \|_2 & = & \sigma_1 & (2.5.8) \end{array}$$

$$\min_{x \neq 0} \frac{\| Ax \|_2}{\| x \|_2} = \sigma_n \qquad (m \ge n)\,. \qquad (2.5.9)$$

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.