NTOU

INTERNATIONAL EDITION

LINEAR ALGEBRA

FOURTH EDITION

STEPHEN H. FRIEDBERG ARNOLD J. INSEL LAWRENCE E. SPENCE

Chen

Acquisitions Editor: *George Lobell*
Editor in Chief: *Sally Yagan*
Production Editor: *Lynn Savino Wendel*
Vice President/Director of Production and Manufacturing: *David W. Riccardi*
Senior Managing Editor: *Linda Mihatov Behrens*
Assistant Managing Editor: *Bayani DeLeon*
Executive Managing Editor: *Kathleen Schiaparelli*
Manufacturing Buyer: *Michael Bell*
Manufacturing Manager: *Trudy Pisciotti*
Editorial Assistant: *Jennifer Brady*
Marketing Manager: *Halee Dinsey*
Marketing Assistant: *Rachel Beckman*
Art Director: *Jayne Conte*
Cover Designer: *Bruce Kenselaar*
Cover Photo Credits: *Anni Albers, Wandbehang We 791 (Orange), 1926/64. Dreifachgewebe: Baumwolle und Kunstseide, schwarz, weiß, Orange 175 × 118 cm. Foto: Gunter Lepkowski, Berlin. Bauhaus-Archiv, Berlin, Inv. Nr. 1575. Lit.: Das Bauhaus webt, Berlin 1998, Nr. 38.*

Prentice Hall

© 2003, 1997, 1989, 1979 by Pearson Education, Inc.
Pearson Education, Inc.
Upper Saddle River, New Jersey 07458

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

This edition may be sold only in those countries to which it is consigned by Pearson Education International. It is not to be re-exported and is not for sale in the U.S.A., Mexico, or Canada.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-120266-9

Pearson Education, Ltd.
Pearson Education Australia Pty. Limited
Pearson Education Singapore, Pte., Ltd
Pearson Education North Asia Ltd
Pearson Education Canada, Ltd.
Pearson Educacion de Mexico, S.A. de C.V.
Pearson Education -- Japan
Pearson Education Malaysia, Pte. Ltd
Pearson Education, Upper Saddle River, New Jersey

The singular value theorem for linear transformations is useful in its matrix form because we can perform numerical computations on matrices. We begin with the definition of the singular values of a matrix.

Definition. *Let A be an $m \times n$ matrix. We define the* **singular values** *of A to be the singular values of the linear transformation L_A.*

Theorem 6.27 (Singular Value Decomposition Theorem for Matrices). *Let A be an $m \times n$ matrix of rank r with the positive singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$, and let Σ be the $m \times n$ matrix defined by*

$$\Sigma_{ij} = \begin{cases} \sigma_i & \text{if } i = j \leq r \\ 0 & \text{otherwise.} \end{cases}$$

Then there exists an $m \times m$ unitary matrix U and an $n \times n$ unitary matrix V such that

$$A = U\Sigma V^*.$$

Proof. Let $T = \mathsf{L}_A \colon \mathsf{F}^n \to \mathsf{F}^m$. By Theorem 6.26, there exist orthonormal bases $\beta = \{v_1, v_2, \ldots, v_n\}$ for F^n and $\gamma = \{u_1, u_2, \ldots, u_m\}$ for F^m such that $\mathsf{T}(v_i) = \sigma_i u_i$ for $1 \leq i \leq r$ and $\mathsf{T}(v_i) = \mathit{0}$ for $i > r$. Let U be the $m \times m$ matrix whose jth column is u_j for all j, and let V be the $n \times n$ matrix whose jth column is v_j for all j. Note that both U and V are unitary matrices.

By Theorem 2.13(a) (p. 90), the jth column of AV is $Av_j = \sigma_j u_j$. Observe that the jth column of Σ is $\sigma_j e_j$, where e_j is the jth standard vector of F^m. So by Theorem 2.13(a) and (b), the jth column of $U\Sigma$ is given by

$$U(\sigma_j e_j) = \sigma_j U(e_j) = \sigma_j u_j.$$

It follows that AV and $U\Sigma$ are $m \times n$ matrices whose corresponding columns are equal, and hence $AV = U\Sigma$. Therefore $A = AVV^* = U\Sigma V^*$. ∎

Definition. *Let A be an $m \times n$ matrix of rank r with positive singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$. A factorization $A = U\Sigma V^*$ where U and V are unitary matrices and Σ is the $m \times n$ matrix defined as in Theorem 6.27 is called a* **singular value decomposition** *of A.*

In the proof of Theorem 6.27, the columns of V are the vectors in β, and the columns of U are the vectors in γ. Furthermore, the nonzero singular values of A are the same as those of L_A; hence they are the square roots of the nonzero eigenvalues of A^*A or of AA^*. (See Exercise 9.)

Example 3

We find a singular value decomposition for $A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix}$.

First observe that for

$$v_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \text{and} \quad v_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix},$$

the set $\beta = \{v_1, v_2, v_3\}$ is an orthonormal basis for R^3 consisting of eigenvectors of A^*A with corresponding eigenvalues $\lambda_1 = 6$, and $\lambda_2 = \lambda_3 = 0$. Consequently, $\sigma_1 = \sqrt{6}$ is the only nonzero singular value of A. Hence, as in the proof of Theorem 6.27, we let V be the matrix whose columns are the vectors in β. Then

$$\Sigma = \begin{pmatrix} \sqrt{6} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}.$$

Also, as in Theorem 6.27, we take

$$u_1 = \frac{1}{\sigma_i} \mathsf{L}_A(v_1) = \frac{1}{\sigma_i} A v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Next choose $u_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, a unit vector orthogonal to u_1, to obtain the orthonormal basis $\gamma = \{u_1, u_2\}$ for R^2, and set

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}.$$

Then $A = U\Sigma V^*$ is the desired singular value decomposition. ◆

The Polar Decomposition of a Square Matrix

A singular value decomposition of a matrix can be used to factor a square matrix in a manner analogous to the factoring of a complex number as the product of a complex number of length 1 and a nonnegative number. In the case of matrices, the complex number of length 1 is replaced by a unitary matrix, and the nonnegative number is replaced by a positive semidefinite matrix.

Theorem 6.28 (Polar Decomposition). *For any square matrix A, there exists a unitary matrix W and a positive semidefinite matrix P such that*

$$A = WP.$$

Furthermore, if A is invertible, then the representation is unique.

Proof. By Theorem 6.27, there exist unitary matrices U and V and a diagonal matrix Σ with nonnegative diagonal entries such that $A = U\Sigma V^*$. So

$$A = U\Sigma V^* = UV^*V\Sigma V^* = WP,$$

where $W = UV^*$ and $P = V\Sigma V^*$. Since W is the product of unitary matrices, W is unitary, and since Σ is positive semidefinite and P is unitarily equivalent to Σ, P is positive semidefinite by Exercise 14 of Section 6.5.

Now suppose that A is invertible and factors as the products

$$A = WP = ZQ,$$

where W and Z are unitary and P and Q are positive semidefinite. Since A is invertible, it follows that P and Q are positive definite and invertible, and therefore $Z^*W = QP^{-1}$. Thus QP^{-1} is unitary, and so

$$I = (QP^{-1})^*(QP^{-1}) = P^{-1}Q^2P^{-1}.$$

Hence $P^2 = Q^2$. Since both P and Q are positive definite, it follows that $P = Q$ by Exercise 17 of Section 6.4. Therefore $W = Z$, and consequently the factorization is unique. ∎

The factorization of a square matrix A as WP where W is unitary and P is positive semidefinite, is called a **polar decomposition** of A.

Example 4

To find the polar decomposition of $A = \begin{pmatrix} 11 & -5 \\ -2 & 10 \end{pmatrix}$, we begin by finding a singular value decomposition $U\Sigma V^*$ of A. The object is to find an orthonormal basis β for R^2 consisting of eigenvectors of A^*A. It can be shown that

$$v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{and} \quad v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

are orthonormal eigenvectors of A^*A with corresponding eigenvalues $\lambda_1 = 200$ and $\lambda_2 = 50$. So $\beta = \{v_1, v_2\}$ is an appropriate basis. Thus $\sigma_1 = \sqrt{200} = 10\sqrt{2}$ and $\sigma_2 = \sqrt{50} = 5\sqrt{2}$ are the singular values of A. So we have

$$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 10\sqrt{2} & 0 \\ 0 & 5\sqrt{2} \end{pmatrix}.$$

Next, we find the columns u_1 and u_2 of U:

$$u_1 = \frac{1}{\sigma_1}Av_1 = \frac{1}{5}\begin{pmatrix} 4 \\ -3 \end{pmatrix} \quad \text{and} \quad u_2 = \frac{1}{\sigma_2}Av_2 = \frac{1}{5}\begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

Thus

$$U = \begin{pmatrix} \frac{4}{5} & \frac{3}{5} \\ -\frac{3}{5} & \frac{4}{5} \end{pmatrix}.$$

Therefore, in the notation of Theorem 6.28, we have

$$W = UV^* = \begin{pmatrix} \frac{4}{5} & \frac{3}{5} \\ -\frac{3}{5} & \frac{4}{5} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{5\sqrt{2}} \begin{pmatrix} 7 & -1 \\ 1 & 7 \end{pmatrix},$$

and

$$P = V\Sigma V^* = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 10\sqrt{2} & 0 \\ 0 & 5\sqrt{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{5}{\sqrt{2}} \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}.$$

◆

The Pseudoinverse

Let V and W be finite-dimensional inner product spaces over the same field, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation. It is desirable to have a linear transformation from W to V that captures some of the essence of an inverse of T even if T is not invertible. A simple approach to this problem is to focus on the "part" of T that is invertible, namely, the restriction of T to $\mathsf{N}(\mathsf{T})^{\perp}$. Let $\mathsf{L}\colon \mathsf{N}(\mathsf{T})^{\perp} \to \mathsf{R}(\mathsf{T})$ be the linear transformation defined by $\mathsf{L}(x) = \mathsf{T}(x)$ for all $x \in \mathsf{N}(\mathsf{T})^{\perp}$. Then L is invertible, and we can use the inverse of L to construct a linear transformation from W to V that salvages some of the benefits of an inverse of T.

Definition. *Let V and W be finite-dimensional inner product spaces over the same field, and let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ be a linear transformation. Let $\mathsf{L}\colon \mathsf{N}(\mathsf{T})^{\perp} \to \mathsf{R}(\mathsf{T})$ be the linear transformation defined by $\mathsf{L}(x) = \mathsf{T}(x)$ for all $x \in \mathsf{N}(\mathsf{T})^{\perp}$. The* **pseudoinverse** *(or Moore-Penrose generalized inverse) of T, denoted by $T^{\dagger}$, is defined as the unique linear transformation from W to V such that*

$$\mathsf{T}^{\dagger}(y) = \begin{cases} \mathsf{L}^{-1}(y) & \text{for } y \in \mathsf{R}(\mathsf{T}) \\ 0 & \text{for } y \in \mathsf{R}(\mathsf{T})^{\perp}. \end{cases}$$

The pseudoinverse of a linear transformation T on a finite-dimensional inner product space exists even if T is not invertible. Furthermore, if T is invertible, then $\mathsf{T}^{\dagger} = \mathsf{T}^{-1}$ because $\mathsf{N}(\mathsf{T})^{\perp} = \mathsf{V}$, and L (as just defined) coincides with T.

As an extreme example, consider the zero transformation $\mathsf{T}_0\colon \mathsf{V} \to \mathsf{W}$ between two finite-dimensional inner product spaces V and W. Then $\mathsf{R}(\mathsf{T}_0) = \{0\}$, and therefore $\mathsf{T}^{\dagger}$ is the zero transformation from W to V.