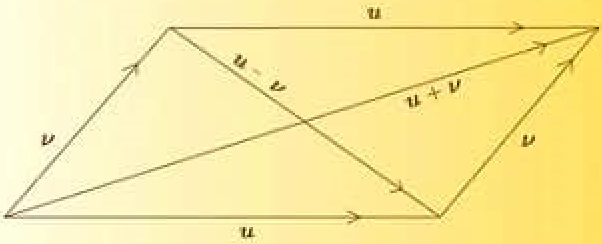
Sheldon Axler

LINEAR ALGEBRA DONE RIGHT

Second Edition



Springer

Undergraduate Texts in Mathematics

Springer
New York
Berlin
Heidelberg
Barcelona
Hong Kong
London
Milan
Paris
Singapore
Tokyo

Undergraduate Texts in Mathematics

Abbott: Understanding Analysis.
Anglin: Mathematics: A Concise History and Philosophy.
Readings in Mathematics.
Anglin/Lambek: The Heritage of Thales.
Readings in Mathematics.
Apostol: Introduction to Analytic Number Theory. Second edition.
Armstrong: Basic Topology.
Armstrong: Groups and Symmetry.
Axler: Linear Algebra Done Right. Second edition.
Beardon: Limits: A New Approach to Real Analysis.
Bak/Newman: Complex Analysis. Second edition.
Banchoff/Wermer: Linear Algebra Through Geometry. Second edition.
Berberian: A First Course in Real Analysis.
Bix: Conics and Cubics: A Concrete Introduction to Algebraic Curves.
Brémaud: An Introduction to Probabilistic Modeling.
Bressoud: Factorization and Primality Testing.
Bressoud: Second Year Calculus.
Readings in Mathematics.
Brickman: Mathematical Introduction to Linear Programming and Game Theory.
Browder: Mathematical Analysis: An Introduction.
Buchmann: Introduction to Cryptography.
Buskes/van Rooij: Topological Spaces: From Distance to Neighborhood.
Callahan: The Geometry of Spacetime: An Introduction to Special and General Relavitity.
Carter/van Brunt: The Lebesgue–Stieltjes Integral: A Practical Introduction.
Cederberg: A Course in Modern Geometries. Second edition.
Childs: A Concrete Introduction to Higher Algebra. Second edition.
Chung: Elementary Probability Theory with Stochastic Processes. Third edition.
Cox/Little/O'Shea: Ideals, Varieties, and Algorithms. Second edition.
Croom: Basic Concepts of Algebraic Topology.
Curtis: Linear Algebra: An Introductory Approach. Fourth edition.
Devlin: The Joy of Sets: Fundamentals of Contemporary Set Theory. Second edition.
Dixmier: General Topology.
Driver: Why Math?
Ebbinghaus/Flum/Thomas: Mathematical Logic. Second edition.
Edgar: Measure, Topology, and Fractal Geometry.
Elaydi: An Introduction to Difference Equations. Second edition.
Exner: An Accompaniment to Higher Mathematics.
Exner: Inside Calculus.
Fine/Rosenberger: The Fundamental Theory of Algebra.
Fischer: Intermediate Real Analysis.
Flanigan/Kazdan: Calculus Two: Linear and Nonlinear Functions. Second edition.
Fleming: Functions of Several Variables. Second edition.
Foulds: Combinatorial Optimization for Undergraduates.
Foulds: Optimization Techniques: An Introduction.
Franklin: Methods of Mathematical Economics.
Frazier: An Introduction to Wavelets Through Linear Algebra.
Gamelin: Complex Analysis.
Gordon: Discrete Probability.
Hairer/Wanner: Analysis by Its History.
Readings in Mathematics.
Halmos: Finite-Dimensional Vector Spaces. Second edition.

(continued after index)

Sheldon Axler

Linear Algebra Done Right

Second Edition

Springer

Sheldon Axler
Mathematics Department
San Francisco State University
San Francisco, CA 94132
USA

Mathematics Subject Classification (1991): 15-01

Library of Congress Cataloging-in-Publication Data
Axler, Sheldon Jay
Linear algebra done right / Sheldon Axler. – 2nd ed.
p. cm. — (Undergraduate texts in mathematics)
Includes index.
ISBN 0-387-98259-0 (alk. paper). – ISBN 0-387-98258-2 (pbk.: alk. paper)
1. Algebra, Linear. I. Title. II. Series.
QA184.A96 1997
512'.5—dc20 97-16664

ISBN 0-387-98259-0 (hardcover) SPIN 10629393
ISBN 0-387-98258-2 (softcover) SPIN 10794473

Springer-Verlag New York Berlin Heidelberg
A member of BertelsmannSpringer Science+Business Media GmbH

Contents

Because $\mathbf{F}^2$ has dimension 2, the last proposition implies that this linearly independent list of length 2 is a basis of $\mathbf{F}^2$ (we do not need to bother checking that it spans $\mathbf{F}^2$).

The next theorem gives a formula for the dimension of the sum of two subspaces of a finite-dimensional vector space.

2.18 Theorem: *If U_1 and U_2 are subspaces of a finite-dimensional vector space, then*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

This formula for the dimension of the sum of two subspaces is analogous to a familiar counting formula: the number of elements in the union of two finite sets equals the number of elements in the first set, plus the number of elements in the second set, minus the number of elements in the intersection of the two sets.

PROOF: Let $(u_1, \ldots, u_m)$ be a basis of $U_1 \cap U_2$; thus $\dim(U_1 \cap U_2) = m$. Because $(u_1, \ldots, u_m)$ is a basis of $U_1 \cap U_2$, it is linearly independent in U_1 and hence can be extended to a basis $(u_1, \ldots, u_m, v_1, \ldots, v_j)$ of U_1 (by 2.12). Thus $\dim U_1 = m + j$. Also extend $(u_1, \ldots, u_m)$ to a basis $(u_1, \ldots, u_m, w_1, \ldots, w_k)$ of U_2; thus $\dim U_2 = m + k$.

We will show that $(u_1, \ldots, u_m, v_1, \ldots, v_j, w_1, \ldots, w_k)$ is a basis of $U_1 + U_2$. This will complete the proof because then we will have

$$\begin{aligned}\dim(U_1 + U_2) &= m + j + k \\ &= (m + j) + (m + k) - m \\ &= \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).\end{aligned}$$

Clearly $\operatorname{span}(u_1, \ldots, u_m, v_1, \ldots, v_j, w_1, \ldots, w_k)$ contains U_1 and U_2 and hence contains $U_1 + U_2$. So to show that this list is a basis of $U_1 + U_2$ we need only show that it is linearly independent. To prove this, suppose

$$a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_j v_j + c_1 w_1 + \cdots + c_k w_k = 0,$$

where all the a's, b's, and c's are scalars. We need to prove that all the a's, b's, and c's equal 0. The equation above can be rewritten as

$$c_1 w_1 + \cdots + c_k w_k = -a_1 u_1 - \cdots - a_m u_m - b_1 v_1 - \cdots - b_j v_j,$$

which shows that $c_1 w_1 + \cdots + c_k w_k \in U_1$. All the w's are in U_2, so this implies that $c_1 w_1 + \cdots + c_k w_k \in U_1 \cap U_2$. Because $(u_1, \ldots, u_m)$ is a basis of $U_1 \cap U_2$, we can write

$$c_1 w_1 + \cdots + c_k w_k = d_1 u_1 + \cdots + d_m u_m$$

Now suppose (c) holds, so that T is surjective. Thus $\operatorname{range} T = V$. From 3.4 we have

$$\begin{aligned}\dim \operatorname{null} T &= \dim V - \dim \operatorname{range} T \\ &= 0,\end{aligned}$$

which implies that $\operatorname{null} T$ equals $\{0\}$. Thus T is injective (by 3.2), and so T is invertible (we already knew that T was surjective). Hence (c) implies (a), completing the proof. ■

and let v be a corresponding nonzero eigenvector. Extend (v) to a basis of V. Then the matrix of T with respect to this basis has the form above. Soon we will see that we can choose a basis of V with respect to which the matrix of T has even more 0's.

The ***diagonal*** of a square matrix consists of the entries along the straight line from the upper left corner to the bottom right corner. For example, the diagonal of the matrix 5.11 consists of the entries $a_{1,1}, a_{2,2}, \dots, a_{n,n}$.

A matrix is called ***upper triangular*** if all the entries below the diagonal equal 0. For example, the 4-by-4 matrix

$$\begin{bmatrix} 6 & 2 & 7 & 5 \\ 0 & 6 & 1 & 3 \\ 0 & 0 & 7 & 9 \\ 0 & 0 & 0 & 8 \end{bmatrix}$$

is upper triangular. Typically we represent an upper-triangular matrix in the form

$$\begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix};$$

the 0 in the matrix above indicates that all entries below the diagonal in this n-by-n matrix equal 0. Upper-triangular matrices can be considered reasonably simple—for n large, an n-by-n upper-triangular matrix has almost half its entries equal to 0.

The following proposition demonstrates a useful connection between upper-triangular matrices and invariant subspaces.

5.12 Proposition: *Suppose $T \in \mathcal{L}(V)$ and $(v_1, \dots, v_n)$ is a basis of V. Then the following are equivalent:*

(a) *the matrix of T with respect to $(v_1, \dots, v_n)$ is upper triangular;*

(b) $Tv_k \in \operatorname{span}(v_1, \dots, v_k)$ *for each $k = 1, \dots, n$;*

(c) $\operatorname{span}(v_1, \dots, v_k)$ *is invariant under T for each $k = 1, \dots, n$.*

Proof: The equivalence of (a) and (b) follows easily from the definitions and a moment's thought. Obviously (c) implies (b). Thus to complete the proof, we need only prove that (b) implies (c). So suppose that (b) holds. Fix $k \in \{1, \dots, n\}$. From (b), we know that

Jordan Form

We know that if V is a complex vector space, then for every $T \in \mathcal{L}(V)$ there is a basis of V with respect to which T has a nice upper-triangular matrix (see 8.28). In this section we will see that we can do even better—there is a basis of V with respect to which the matrix of T contains zeros everywhere except possibly on the diagonal and the line directly above the diagonal.

We begin by describing the nilpotent operators. Consider, for example, the nilpotent operator $N \in \mathcal{L}(\mathbf{F}^n)$ defined by

$$N(z_1, \dots, z_n) = (0, z_1, \dots, z_{n-1}).$$

If $v = (1, 0, \dots, 0)$, then clearly $(v, Nv, \dots, N^{n-1}v)$ is a basis of $\mathbf{F}^n$ and $(N^{n-1}v)$ is a basis of null N, which has dimension 1.

As another example, consider the nilpotent operator $N \in \mathcal{L}(\mathbf{F}^5)$ defined by

8.39 $$N(z_1, z_2, z_3, z_4, z_5) = (0, z_1, z_2, 0, z_4).$$

Unlike the nilpotent operator discussed in the previous paragraph, for this nilpotent operator there does not exist a vector $v \in \mathbf{F}^5$ such that $(v, Nv, N^2v, N^3v, N^4v)$ is a basis of $\mathbf{F}^5$. However, if $v_1 = (1, 0, 0, 0, 0)$ and $v_2 = (0, 0, 0, 1, 0)$, then $(v_1, Nv_1, N^2v_1, v_2, Nv_2)$ is a basis of $\mathbf{F}^5$ and (N^2v_1, Nv_2) is a basis of null N, which has dimension 2.

Suppose $N \in \mathcal{L}(V)$ is nilpotent. For each nonzero vector $v \in V$, let $m(v)$ denote the largest nonnegative integer such that $N^{m(v)}v \neq 0$. For example, if $N \in \mathcal{L}(\mathbf{F}^5)$ is defined by 8.39, then $m(1, 0, 0, 0, 0) = 2$.

Obviously $m(v)$ depends on N as well as on v, but the choice of N will be clear from the context.

The lemma below shows that every nilpotent operator $N \in \mathcal{L}(V)$ behaves similarly to the example defined by 8.39, in the sense that there is a finite collection of vectors $v_1, \dots, v_k \in V$ such that the nonzero vectors of the form $N^j v_r$ form a basis of V; here r varies from 1 to k and j varies from 0 to $m(v_r)$.

8.40 Lemma: *If $N \in \mathcal{L}(V)$ is nilpotent, then there exist vectors $v_1, \dots, v_k \in V$ such that*

(a) $(v_1, Nv_1, \dots, N^{m(v_1)}v_1, \dots, v_k, Nv_k, \dots, N^{m(v_k)}v_k)$ *is a basis of V;*

(b) $(N^{m(v_1)}v_1, \dots, N^{m(v_k)}v_k)$ *is a basis of* null N.