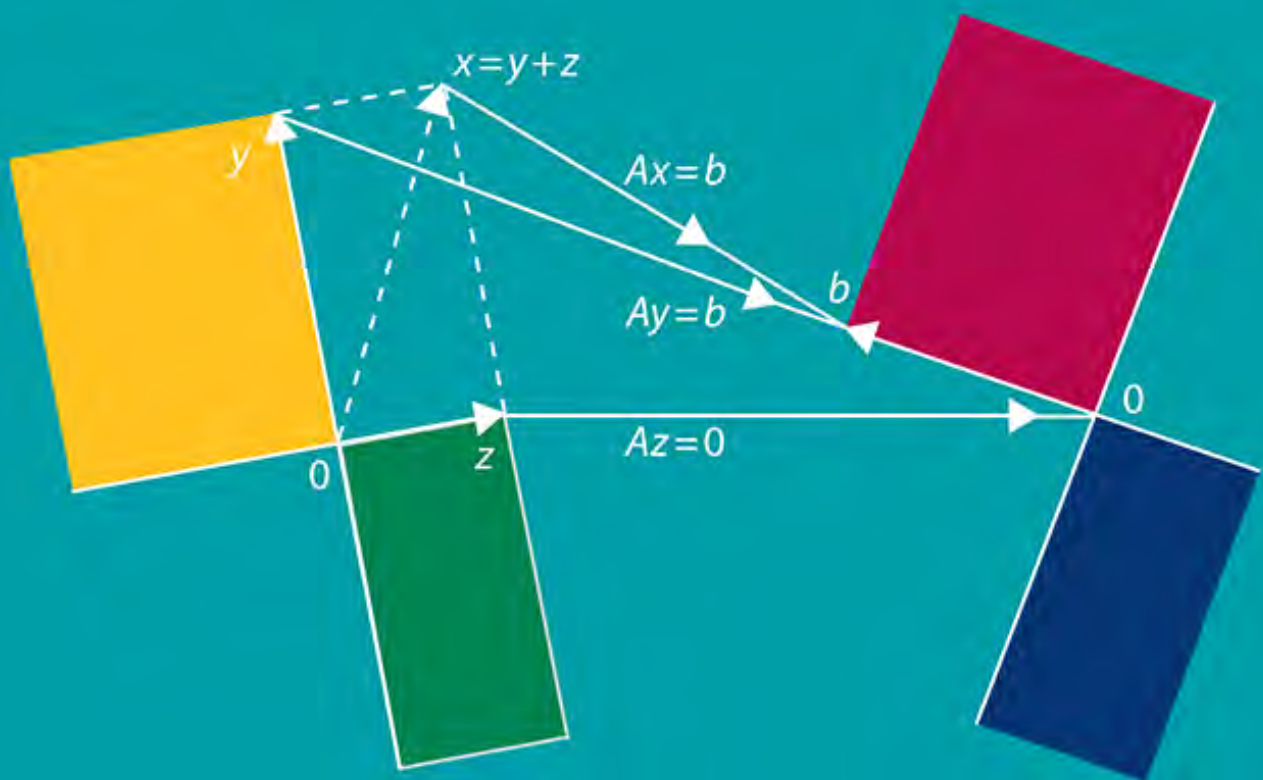


Introduction to

LINEAR ALGEBRA

FIFTH EDITION



GILBERT STRANG

INTRODUCTION TO LINEAR ALGEBRA

Fifth Edition

GILBERT STRANG

Massachusetts Institute of Technology

WELLESLEY - CAMBRIDGE PRESS
Box 812060 Wellesley MA 02482

Introduction to Linear Algebra, 5th Edition

Copyright ©2016 by Gilbert Strang

ISBN 978-0-9802327-7-6

All rights reserved. No part of this book may be reproduced or stored or transmitted by any means, including photocopying, without written permission from Wellesley - Cambridge Press. Translation in any language is strictly prohibited — authorized translations are arranged by the publisher.

L^AT_EX typesetting by Ashley C. Fernandes (info@problemsolvingpathway.com)

Printed in the United States of America

9 8 7 6 5 4 3

QA184.S78 2016 512'.5 93-14092

Other texts from Wellesley - Cambridge Press

Computational Science and Engineering , Gilbert Strang	ISBN 978-0-9614088-1-7
Wavelets and Filter Banks , Gilbert Strang and Truong Nguyen	ISBN 978-0-9614088-7-9
Introduction to Applied Mathematics , Gilbert Strang	ISBN 978-0-9614088-0-0
Calculus Third Edition (2017) , Gilbert Strang	ISBN 978-0-9802327-5-2
Algorithms for Global Positioning , Kai Borre & Gilbert Strang	ISBN 978-0-9802327-3-8
Essays in Linear Algebra , Gilbert Strang	ISBN 978-0-9802327-6-9
Differential Equations and Linear Algebra , Gilbert Strang	ISBN 978-0-9802327-9-0
An Analysis of the Finite Element Method , 2008 edition, Gilbert Strang and George Fix	ISBN 978-0-9802327-0-7

Wellesley - Cambridge Press

Box 812060

Wellesley MA 02482 USA

www.wellesleycambridge.com

linearalgebrabook@gmail.com

math.mit.edu/~gs

phone (781) 431-8488

fax (617) 253-4358

The website for this book is math.mit.edu/linearalgebra.

The Solution Manual can be printed from that website.

Course material including syllabus and exams and also videotaped lectures are available on the book website and the teaching website: web.mit.edu/18.06

Linear Algebra is included in MIT's OpenCourseWare site ocw.mit.edu.

This provides video lectures of the full linear algebra course 18.06 and 18.06 SC.

MATLAB® is a registered trademark of The MathWorks, Inc.

The front cover captures a central idea of linear algebra.

$Ax = b$ is solvable when b is in the (red) column space of A .

One particular solution y is in the (yellow) row space: $Ay = b$.

Add any vector z from the (green) nullspace of A : $Az = 0$.

The complete solution is $x = y + z$. Then $Ax = Ay + Az = b$.

The cover design was the inspiration of Lois Sellers and Gail Corbett.

Table of Contents

1	Introduction to Vectors	1
1.1	Vectors and Linear Combinations	2
1.2	Lengths and Dot Products	11
1.3	Matrices	22
2	Solving Linear Equations	31
2.1	Vectors and Linear Equations	31
2.2	The Idea of Elimination	46
2.3	Elimination Using Matrices	58
2.4	Rules for Matrix Operations	70
2.5	Inverse Matrices	83
2.6	Elimination = Factorization: $A = LU$	97
2.7	Transposes and Permutations	109
3	Vector Spaces and Subspaces	123
3.1	Spaces of Vectors	123
3.2	The Nullspace of A : Solving $Ax = 0$ and $Rx = 0$	135
3.3	The Complete Solution to $Ax = b$	150
3.4	Independence, Basis and Dimension	164
3.5	Dimensions of the Four Subspaces	181
4	Orthogonality	194
4.1	Orthogonality of the Four Subspaces	194
4.2	Projections	206
4.3	Least Squares Approximations	219
4.4	Orthonormal Bases and Gram-Schmidt	233
5	Determinants	247
5.1	The Properties of Determinants	247
5.2	Permutations and Cofactors	258
5.3	Cramer's Rule, Inverses, and Volumes	273

6 Eigenvalues and Eigenvectors	288
6.1 Introduction to Eigenvalues	288
6.2 Diagonalizing a Matrix	304
6.3 Systems of Differential Equations	319
6.4 Symmetric Matrices	338
6.5 Positive Definite Matrices	350
7 The Singular Value Decomposition (SVD)	364
7.1 Image Processing by Linear Algebra	364
7.2 Bases and Matrices in the SVD	371
7.3 Principal Component Analysis (PCA by the SVD)	382
7.4 The Geometry of the SVD	392
8 Linear Transformations	401
8.1 The Idea of a Linear Transformation	401
8.2 The Matrix of a Linear Transformation	411
8.3 The Search for a Good Basis	421
9 Complex Vectors and Matrices	430
9.1 Complex Numbers	431
9.2 Hermitian and Unitary Matrices	438
9.3 The Fast Fourier Transform	445
10 Applications	452
10.1 Graphs and Networks	452
10.2 Matrices in Engineering	462
10.3 Markov Matrices, Population, and Economics	474
10.4 Linear Programming	483
10.5 Fourier Series: Linear Algebra for Functions	490
10.6 Computer Graphics	496
10.7 Linear Algebra for Cryptography	502
11 Numerical Linear Algebra	508
11.1 Gaussian Elimination in Practice	508
11.2 Norms and Condition Numbers	518
11.3 Iterative Methods and Preconditioners	524
12 Linear Algebra in Probability & Statistics	535
12.1 Mean, Variance, and Probability	535
12.2 Covariance Matrices and Joint Probabilities	546
12.3 Multivariate Gaussian and Weighted Least Squares	555
Matrix Factorizations	563
Index	565
Six Great Theorems / Linear Algebra in a Nutshell	574

Preface

I am happy for you to see this Fifth Edition of Introduction to Linear Algebra. This is the text for my video lectures on MIT's OpenCourseWare (ocw.mit.edu and also **YouTube**). I hope those lectures will be useful to you (maybe even enjoyable!).

Hundreds of colleges and universities have chosen this textbook for their basic linear algebra course. A sabbatical gave me a chance to prepare two new chapters about probability and statistics and understanding data. Thousands of other improvements too—probably only noticed by the author... Here is a new addition for students and all readers:

Every section opens with a brief summary to explain its contents. When you read a new section, and when you revisit a section to review and organize it in your mind, those lines are a quick guide and an aid to memory.

Another big change comes on this book's website math.mit.edu/linearalgebra. That site now contains solutions to the Problem Sets in the book. With unlimited space, this is much more flexible than printing short solutions. There are three key websites:

ocw.mit.edu Messages come from thousands of students and faculty about linear algebra on this OpenCourseWare site. The 18.06 and 18.06 SC courses include video lectures of a complete semester of classes. Those lectures offer an independent review of the whole subject based on this textbook—the professor's time stays free and the student's time can be 2 a.m. (The reader doesn't have to be in a class at all.) Six million viewers around the world have seen these videos (*amazing*). I hope you find them helpful.

web.mit.edu/18.06 This site has homeworks and exams (with solutions) for the current course as it is taught, and as far back as 1996. There are also review questions, Java demos, Teaching Codes, and short essays (*and the video lectures*). My goal is to make this book as useful to you as possible, with all the course material we can provide.

math.mit.edu/linearalgebra This has become an active website. It now has Solutions to Exercises—with space to explain ideas. There are also new exercises from many different sources—practice problems, development of textbook examples, codes in MATLAB and *Julia* and *Python*, plus whole collections of exams (18.06 and others) for review.

Please visit this linear algebra site. Send suggestions to linearalgebrabook@gmail.com

The Fifth Edition

The cover shows the **Four Fundamental Subspaces**—the row space and nullspace are on the left side, the column space and the nullspace of A^T are on the right. It is not usual to put the central ideas of the subject on display like this! When you meet those four spaces in Chapter 3, you will understand why that picture is so central to linear algebra.

Those were named the Four Fundamental Subspaces in my first book, and they start from a matrix A . Each row of A is a vector in n -dimensional space. When the matrix has m rows, each column is a vector in m -dimensional space. The crucial operation in linear algebra is to take **linear combinations of column vectors**. This is exactly the result of a matrix-vector multiplication. Ax is a combination of the columns of A .

When we take *all* combinations Ax of the column vectors, we get the *column space*. If this space includes the vector b , we can solve the equation $Ax = b$.

May I call special attention to Section 1.3, where these ideas come early—with two specific examples. You are not expected to catch every detail of vector spaces in one day! But you will see the first matrices in the book, and a picture of their column spaces. There is even an *inverse matrix* and its connection to calculus. You will be learning the language of linear algebra in the best and most efficient way: by using it.

Every section of the basic course ends with a large collection of review problems. They ask you to use the ideas in that section—the dimension of the column space, a basis for that space, the rank and inverse and determinant and eigenvalues of A . Many problems look for computations by hand on a small matrix, and they have been highly praised. The *Challenge Problems* go a step further, and sometimes deeper. Let me give four examples:

Section 2.1: Which row exchanges of a Sudoku matrix produce another Sudoku matrix?

Section 2.7: If P is a permutation matrix, why is some power P^k equal to I ?

Section 3.4: If $Ax = b$ and $Cx = b$ have the same solutions for every b , does A equal C ?

Section 4.1: What conditions on the four vectors r , n , c , ℓ allow them to be bases for the row space, the nullspace, the column space, and the left nullspace of a 2 by 2 matrix?

The Start of the Course

The equation $Ax = b$ uses the language of linear combinations right away. The vector Ax is a combination of the columns of A . The equation is asking for a combination that produces b . The solution vector x comes at three levels and all are important:

1. **Direct solution** to find x by forward elimination and back substitution.
2. **Matrix solution** using the inverse matrix: $x = A^{-1}b$ (if A has an inverse).
3. **Particular solution** (to $Ay = b$) plus **nullspace solution** (to $Az = 0$).

That vector space solution $x = y + z$ is shown on the cover of the book.

Direct elimination is the most frequently used algorithm in scientific computing. The matrix A becomes triangular—then solutions come quickly. We also see bases for the four subspaces. But don't spend forever on practicing elimination ... good ideas are coming.

The speed of every new supercomputer is tested on $Ax = b$: pure linear algebra. But even a supercomputer doesn't want the inverse matrix: *too slow*. Inverses give the simplest formula $x = A^{-1}b$ but not the top speed. And everyone must know that determinants are even slower—there is no way a linear algebra course should begin with formulas for the determinant of an n by n matrix. Those formulas have a place, but not first place.

Structure of the Textbook

Already in this preface, you can see the style of the book and its goal. That goal is serious, to explain this beautiful and useful part of mathematics. You will see how the applications of linear algebra reinforce the key ideas. This book moves gradually and steadily from *numbers* to *vectors* to *subspaces*—each level comes naturally and everyone can get it.

Here are 12 points about learning and teaching from this book:

1. Chapter 1 starts with vectors and dot products. If the class has met them before, focus quickly on linear combinations. Section 1.3 provides three independent vectors whose combinations fill all of 3-dimensional space, and three dependent vectors in a plane. *Those two examples are the beginning of linear algebra.*
2. Chapter 2 shows the row picture and the column picture of $Ax = b$. The heart of linear algebra is in that connection between the rows of A and the columns of A : the same numbers but very different pictures. Then begins the algebra of matrices: an elimination matrix E multiplies A to produce a zero. The goal is to capture the whole process—start with A , multiply by E 's, end with U .

Elimination is seen in the beautiful form $A = LU$. The **lower triangular** L holds the forward elimination steps, and U is **upper triangular** for back substitution.

3. Chapter 3 is linear algebra at the best level: **subspaces**. The column space contains all linear combinations of the columns. The crucial question is: *How many of those columns are needed?* The answer tells us the dimension of the column space, and the key information about A . We reach the Fundamental Theorem of Linear Algebra.
4. With more equations than unknowns, it is almost sure that $Ax = b$ has no solution. We cannot throw out every measurement that is close but not perfectly exact! When we solve by **least squares**, the key will be the matrix $A^T A$. This wonderful matrix appears everywhere in applied mathematics, when A is rectangular.
5. **Determinants** give formulas for all that has come before—Cramer's Rule, inverse matrices, volumes in n dimensions. We don't need those formulas to compute. They slow us down. But $\det A = 0$ tells when a matrix is singular: this is the key to eigenvalues.

6. **Section 6.1 explains eigenvalues for 2 by 2 matrices.** Many courses want to see eigenvalues early. It is completely reasonable to come here directly from Chapter 3, because the determinant is easy for a 2 by 2 matrix. *The key equation is $Ax = \lambda x$.*
 Eigenvalues and eigenvectors are an astonishing way to understand a square matrix. They are not for $Ax = b$, they are for dynamic equations like $du/dt = Au$. The idea is always the same: *follow the eigenvectors*. In those special directions, A acts like a single number (the eigenvalue λ) and the problem is one-dimensional. An essential highlight of Chapter 6 is **diagonalizing a symmetric matrix**. When all the eigenvalues are positive, the matrix is “positive definite”. This key idea connects the whole course—positive pivots and determinants and eigenvalues and energy. I work hard to reach this point in the book and to explain it by examples.
7. Chapter 7 is new. It introduces **singular values** and **singular vectors**. They separate all matrices into simple pieces, ranked in order of their importance. You will see one way to compress an image. Especially you can analyze a matrix full of data.
8. Chapter 8 explains **linear transformations**. This is geometry without axes, algebra with no coordinates. When we choose a basis, we reach the best possible matrix.
9. Chapter 9 moves from real numbers and vectors to complex vectors and matrices. The Fourier matrix F is the most important complex matrix we will ever see. And the **Fast Fourier Transform** (multiplying quickly by F and F^{-1}) is revolutionary.
10. Chapter 10 is full of applications, more than any single course could need:
 - 10.1 *Graphs and Networks*—leading to the edge-node matrix for Kirchhoff’s Laws
 - 10.2 *Matrices in Engineering*—differential equations parallel to matrix equations
 - 10.3 *Markov Matrices*—as in Google’s *PageRank* algorithm
 - 10.4 *Linear Programming*—a new requirement $x \geq 0$ and minimization of the cost
 - 10.5 *Fourier Series*—linear algebra for functions and digital signal processing
 - 10.6 *Computer Graphics*—matrices move and rotate and compress images
 - 10.7 *Linear Algebra in Cryptography*—this new section was fun to write. The Hill Cipher is not too secure. It uses modular arithmetic: integers from 0 to $p - 1$. Multiplication gives $4 \times 5 \equiv 1 \pmod{19}$. For decoding this gives $4^{-1} \equiv 5$.
11. How should computing be included in a linear algebra course? It can open a new understanding of matrices—every class will find a balance. **MATLAB** and *Maple* and *Mathematica* are powerful in different ways. *Julia* and *Python* are free and directly accessible on the Web. Those newer languages are powerful too!
 Basic commands begin in Chapter 2. Then Chapter 11 moves toward professional algorithms. You can upload and download codes for this course on the website.
12. Chapter 12 on Probability and Statistics is new, with truly important applications. When random variables are not independent we get covariance matrices. Fortunately they are symmetric positive definite. The linear algebra in Chapter 6 is needed now.

The Variety of Linear Algebra

Calculus is mostly about one special operation (the derivative) and its inverse (the integral). Of course I admit that calculus could be important But so many applications of mathematics are discrete rather than continuous, digital rather than analog. The century of data has begun! You will find a light-hearted essay called “Too Much Calculus” on my website. *The truth is that vectors and matrices have become the language to know.*

Part of that language is the wonderful variety of matrices. Let me give three examples:

<i>Symmetric matrix</i>	<i>Orthogonal matrix</i>	<i>Triangular matrix</i>
$\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$	$\frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

A key goal is learning to “read” a matrix. You need to see the meaning in the numbers. This is really the essence of mathematics—patterns and their meaning.

I have used *italics* and **boldface** to pick out the key words on each page. I know there are times when you want to read quickly, looking for the important lines.

May I end with this thought for professors. You might feel that the direction is right, and wonder if your students are ready. **Just give them a chance!** Literally thousands of students have written to me, frequently with suggestions and surprisingly often with thanks. They know this course has a purpose, because the professor and the book are on their side. Linear algebra is a fantastic subject, enjoy it.

Help With This Book

The greatest encouragement of all is the feeling that you are doing something worthwhile with your life. Hundreds of generous readers have sent ideas and examples and corrections (and favorite matrices) that appear in this book. *Thank you all.*

One person has helped with every word in this book. He is Ashley C. Fernandes, who prepared the L^AT_EX files. It is now six books that he has allowed me to write and rewrite, aiming for accuracy and also for life. Working with friends is a happy way to live.

Friends inside and outside the MIT math department have been wonderful. Alan Edelman for *Julia* and much more, Alex Townsend for the flag examples in 7.1, and Peter Kempthorne for the finance example in 7.3: those stand out. Don Spickler’s website on cryptography is simply excellent. I thank Jon Bloom, Jack Dongarra, Hilary Finucane, Pavel Grinfeld, Randy LeVeque, David Vogan, Liang Wang, and Karen Willcox. The “eigenfaces” in 7.3 came from Matthew Turk and Jeff Jauregui. And the big step to singular values was accelerated by Raj Rao’s great course at Michigan.

This book owes so much to my happy sabbatical in Oxford. Thank you, Nick Trefethen and everyone. Especially you the reader! Best wishes in your work.

Background of the Author

This is my 9th textbook on linear algebra, and I hesitate to write about myself. It is the mathematics that is important, and the reader. The next paragraphs add something brief and personal, as a way to say that textbooks are written by people.

I was born in Chicago and went to school in Washington and Cincinnati and St. Louis. My college was MIT (and my linear algebra course was *extremely abstract*). After that came Oxford and UCLA, then back to MIT for a very long time. I don't know how many thousands of students have taken 18.06 (more than 6 million when you include the videos on ocw.mit.edu). The time for a fresh approach was right, because this fantastic subject was only revealed to math majors—we needed to open linear algebra to the world.

I am so grateful for a life of teaching mathematics, more than I could possibly tell you.

Gilbert Strang

PS I hope the next book (2018 ?) will include *Learning from Data*. This subject is growing quickly, especially “deep learning”. By knowing a function on a training set of old data, we approximate the function on new data. The approximation only uses one simple non-linear function $f(x) = \max(0, x)$. It is n matrix multiplications that we optimize to make the learning deep: $x_1 = f(A_1x + b_1)$, $x_2 = f(A_2x_1 + b_2)$, \dots , $x_n = f(A_nx_{n-1} + b_n)$. Those are $n - 1$ hidden layers between the input x and the output x_n —which approximates $F(x)$ on the training set.

THE MATRIX ALPHABET

A	Any Matrix	P	Permutation Matrix
B	Basis Matrix	P	Projection Matrix
C	Cofactor Matrix	Q	Orthogonal Matrix
D	Diagonal Matrix	R	Upper Triangular Matrix
E	Elimination Matrix	R	Reduced Echelon Matrix
F	Fourier Matrix	S	Symmetric Matrix
H	Hadamard Matrix	T	Linear Transformation
I	Identity Matrix	U	Upper Triangular Matrix
J	Jordan Matrix	U	Left Singular Vectors
K	Stiffness Matrix	V	Right Singular Vectors
L	Lower Triangular Matrix	X	Eigenvector Matrix
M	Markov Matrix	Λ	Eigenvalue Matrix
N	Nullspace Matrix	Σ	Singular Value Matrix