INTRODUCTION TO LINEAR ALGEBRA

Indian Edition

GILBERT STRANG

Massachusetts Institute of Technology

WELLESLEY PUBLISHERS - INDIA

Introduction to Linear Algebra, Indian Edition

ISBN 978-1-7331466-1-6

LaTeX typesetting by Ashley C. Fernandes (info@problemsolvingpathway.com)

Wellesley Publishers, India	**www.wellesleypublishers.com/buy.html**
Gilbert Strang@MIT	**math.mit.edu/∼gs**

The website for this book is **math.mit.edu/linearalgebra**
That site will link to course material and video lectures on YouTube and OCW

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

The front cover captures a central idea of linear algebra.
$A\boldsymbol{x} = \boldsymbol{b}$ is solvable when $\boldsymbol{b}$ is in the (orange) column space of A.
One particular solution $\boldsymbol{y}$ is in the (red) row space: $A\boldsymbol{y} = \boldsymbol{b}$.
Add any vector $\boldsymbol{z}$ from the (green) nullspace of A: $A\boldsymbol{z} = \boldsymbol{0}$.
The complete solution is $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$. Then $A\boldsymbol{x} = A\boldsymbol{y} + A\boldsymbol{z} = \boldsymbol{b}$.

Table of Contents

Preface

I will be happy with this preface if three important points come through clearly :

1. The beauty and variety of linear algebra, and its extreme usefulness
2. The goals of this book, and the new features in this Indian Edition
3. The steady support from our linear algebra websites and the video lectures

May I begin with notes about two websites that are constantly used.

ocw.mit.edu Messages come from thousands of students and faculty about linear algebra on MIT's OpenCourseWare site. The Math 18.06 course includes video lectures of a complete semester of classes and 18.06SC adds problem solutions by MIT teachers. The video lectures offer an independent review of the whole subject based on this book. Ten million viewers around the world have seen these videos (*amazing*). I hope you find them helpful.

math.mit.edu/linearalgebra This website is a permanent record of ideas and good problems. Solutions are now included for this Indian Edition. Several sections of the book are directly available online, plus notes on teaching linear algebra. The content is growing quickly and contributions are welcome from everyone.

The Indian Edition

A new set of video lectures was added to ocw.mit.edu in 2019 for Math 18.065 : Linear Algebra and Learning from Data. That topic is now introduced at the end of this Indian Edition. Readers want to know about "deep learning" and we have brought pages from our new book into this book.

The book 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! You will meet those four spaces in Chapter 3, and you will understand why that picture is 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 taking ***linear combinations of vectors***. (That idea starts on page 1 of the book and never stops.) *When we take all linear combinations 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 attention to Section 1.3 in which 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, and even an *inverse matrix*. 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 now ends with ***Challenge Problems***. They follow a large collection of review problems, which 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 new Challenge Problems go a step further, and sometimes they go deeper. Let me give four examples:

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

Section 2.4: From the shapes of A, B, C, is it faster to compute AB times C or A times BC?

Background: Multiplying *AB times C gives the same answer as A times BC*. This simple statement is the reason behind the rule for matrix multiplication. If AB is a big matrix and C is a vector, it's faster to multiply BC first.

That decision is the key to **backpropagation**—the central algorithm of deep learning. We are computing derivatives by the chain rule, to optimize the weights. Going backwards from the output can make the computations a thousand times faster.

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

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 of A: $x = A^{-1}b$ (if A has an inverse).

3. ***Vector space solution*** $x = y + z$ as shown on the cover of the book:

   Particular solution (to $Ay = b$) plus ***nullspace solution*** (to $Az = 0$)

Direct elimination is the most frequently used algorithm in scientific computing, and the idea is not hard. Simplify the matrix A so it becomes triangular—then all solutions come quickly. I don't spend forever on practicing elimination, it will get learned.

The speed of every new supercomputer is tested on $Ax = b$: it's pure linear algebra. **Top500.org** is a list of the world's fastest computers. In 2008 they reached *petaflop speed* (10^{15} operations per second) by solving many equations in parallel. For high performance, computers avoid operating on single numbers, they feed on whole submatrices.

The processors in the Roadrunner are based on the Cell Engine in PlayStation 3. What can I say, video games are now the largest market for the fastest computations.

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. I hope every teacher will learn something new; familiar ideas can be seen in a new way. The book moves gradually and steadily from *numbers* to *vectors* to *subspaces*—each level comes naturally and everyone can get it.

Here are ten points about the organization of the ten chapters :

1. Chapter 1 starts with vectors and dot products. 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: 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 here is to capture the whole process—start with A and end with an ***upper triangular*** U.

   Elimination is seen in the beautiful form $\boldsymbol{A} = \boldsymbol{LU}$. The ***lower triangular*** L holds all the forward elimination steps, and U is the matrix 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. Chapter 4 has m equations and only n unknowns. It is almost sure that $Ax = b$ has no solution. We cannot throw out equations that are close but not perfectly exact. When we solve by ***least squares***, the key will be the matrix $A^{\mathrm{T}}A$. This wonderful matrix $A^{\mathrm{T}}A$ appears everywhere in applied mathematics, when A is rectangular.

5. ***Determinants*** in Chapter 5 give formulas for all that has come before—inverses, pivots, volumes in n-dimensional space, and more. We don't need those formulas to compute! They slow us down. But $\det A = 0$ tells when a matrix is singular, and that test is the key to eigenvalues.

6. ***Section 6.1 introduces 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.

One highlight is ***diagonalizing a symmetric matrix***. Another highlight—more important every day—is the diagonalization of *any matrix*. This needs two sets of eigenvectors, not one, and they come (of course!) from $A^{\mathrm{T}}A$ and AA^{T}. This **Singular Value Decomposition** often marks the end of the basic course.

Every section in the basic course ends with a ***Review of the Key Ideas***.

7. Chapter 7 explains the ***linear transformation*** approach—it is linear algebra without coordinates, the ideas without computations. Chapter 9 is the opposite—all about how $Ax = b$ and $Ax = \lambda x$ are really solved.

8. Chapter 8 is full of applications, more than any single course could need:

8.1 *Matrices in Engineering*—differential equations replaced by matrix equations

8.2 *Graphs and Networks*—leading to the edge-node matrix for Kirchhoff's Laws

8.3 *Markov Matrices*—as in Google's *PageRank* algorithm

8.4 *Linear Programming*—a new requirement $x \geq 0$ and minimization of the cost

8.5 *Fourier Series*—linear algebra for functions and digital signal processing

8.6 *Matrices in Statistics and Probability*—$Ax = b$ is weighted by average errors

8.7 *Computer Graphics*—matrices move and rotate and compress images.

9. How should computing be included in a linear algebra course? It can open a new understanding of matrices—every class will find a balance. I chose the language of MATLAB as a direct way to describe linear algebra: eig(ones(4)) will produce the eigenvalues $4, 0, 0, 0$ of the 4 by 4 all-ones matrix. *Go to **netlib.org** for codes.* You can freely choose a different system. More and more software is open source.

10. Chapter 10 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 a revolutionary algorithm.

The website ***math.mit.edu/linearalgebra*** provides ideas about teaching and learning.

This Indian Edition includes four special appendices. Singular vectors lead to *Principal Component Analysis* (to find the important part of a matrix). **Neural nets lead to "deep learning"**. This idea is changing our world: Machine learning is based on linear algebra. It is explained in the book **Linear Algebra and Learning from Data**.

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:

Symmetric matrix $\begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$ ***Orthogonal matrix*** $\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$ ***Triangular matrix*** $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

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.

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

For many years, the preparation of my textbooks has depended on Ashley C. Fernandes. I scan handwritten pages to Mumbai. Ashley returns them in the perfect form that you see here. Then I work more on the text, trying to make the ideas clear. After many iterations the book is ready to print—and Ashley has made it beautiful.

Background of the Author

This is my tenth 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 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 ten 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.**

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.*

Linear Algebra and Learning from Data This is the new textbook for the applied linear algebra course 18.065 at MIT. It starts with the basic factorizations of a matrix:

$$\boldsymbol{A = LU} \quad \boldsymbol{A = QR} \quad \boldsymbol{A = X\Lambda X^{-1}} \quad \boldsymbol{S = Q\Lambda Q^{\mathrm{T}}} \quad \boldsymbol{A = U\Sigma V^{\mathrm{T}}} \quad \boldsymbol{A = CMR}$$

The last three are fundamental for data analysis—eigenvalues Λ of a symmetric matrix S, singular values Σ of any matrix A (the SVD is a highlight of this book too), and column-mixing-row factors $A = CMR$. **See the Math 18.065 videos on ocw.mit.edu** *The goal of deep learning is to find patterns in the training data*. Matrix multiplication is interwoven with the nonlinear ramp function ReLU $(x) = \textbf{max}\,(\boldsymbol{0}, \boldsymbol{x})$. The result is a learning function that can interpret *new data*. The textbook explains how and why this succeeds—even in the classroom. Linear algebra and student projects are the keys.

THE MATRIX ALPHABET

$\boldsymbol{A}$	Any Matrix	$\boldsymbol{R}$	Reduced Echelon Matrix
$\boldsymbol{D}$	Diagonal Matrix	$\boldsymbol{R}$	Upper Triangular Matrix
$\boldsymbol{E}$	Elimination Matrix	$\boldsymbol{S}$	Symmetric Matrix
$\boldsymbol{F}$	Fourier Matrix	$\boldsymbol{S}$	Sample Covariance Matrix
$\boldsymbol{I}$	Identity Matrix	$\boldsymbol{T}$	Linear Transformation
$\boldsymbol{J}$	Jordan Matrix	$\boldsymbol{U}$	Upper Triangular Matrix
$\boldsymbol{L}$	Lower Triangular Matrix	$\boldsymbol{U}$	Left Singular Vectors
$\boldsymbol{M}$	Markov Matrix	$\boldsymbol{V}$	Right Singular Vectors
$\boldsymbol{P}$	Permutation Matrix	$\boldsymbol{X}$	Eigenvector Matrix
$\boldsymbol{P}$	Projection Matrix	$\boldsymbol{\Lambda}$	Eigenvalue Matrix
$\boldsymbol{Q}$	Orthogonal Matrix	$\boldsymbol{\Sigma}$	Singular Value Matrix

Video lectures: OpenCourseWare ocw.mit.edu and YouTube (**Math 18.06 and 18.065**)
Book websites: **math.mit.edu/linearalgebra** and **math.mit.edu/learningfromdata**