Jonathan S. Golan

The Linear Algebra a Beginning Graduate Student Ought to Know

Second Edition

线性代数 第2版

Springer

世界图书出版公司
www.wpcbj.com.cn

The Linear Algebra a Beginning Graduate Student Ought to Know

Second Edition

by

JONATHAN S. GOLAN

University of Haifa, Israel

Springer

图书在版编目(CIP)数据

线性代数 = The Linear Algebra A Beginning Graduate Student ought to Know:英文/(以)戈兰(Golan,J. S.)著. —影印本. —北京:世界图书出版公司北京公司,2011. 7

ISBN 978 - 7 - 5100 - 3745 - 0

Ⅰ. ①线… Ⅱ. ①戈… Ⅲ. ①线性代数—高等学校—教材—英文 Ⅳ. ①O151. 2

中国版本图书馆 CIP 数据核字(2011)第 138935 号

书　　名: The Linear Algebra a Beginning Graduate Student Ought to Know 2nd ed.
作　　者: Jonathan S. Golan
中 译 名: 线性代数 第 2 版
责任编辑: 高蓉　刘慧

出 版 者: 世界图书出版公司北京公司
印 刷 者: 三河市国英印务有限公司
发　　行: 世界图书出版公司北京公司(北京朝内大街 137 号　100010)
联系电话: 010 - 64021602,010 - 64015659
电子信箱: kjb@ wpcbj. com. cn

开　　本: 24 开
印　　张: 19
版　　次: 2011 年 07 月
版权登记: 图字:01 - 2011 - 3247

书　　号: 978 - 7 - 5100 - 3745 - 0/O · 899　　**定　　价:** 49. 00 元

THE LINEAR ALGEBRA A BEGINNING GRADUATE STUDENT OUGHT TO KNOW

To my grandsons: Shachar, Eitan, and Sarel

תן לחכם ויחכם עוד

(משלי, פרק ט')

For whom is this book written?

Crow's Law: Do not think what you want to think until you know what you ought to know.[1]

Linear algebra is a living, active branch of mathematical research which is central to almost all other areas of mathematics and which has important applications in all branches of the physical and social sciences and in engineering. However, in recent years the content of linear algebra courses required to complete an undergraduate degree in mathematics – and even more so in other areas – at all but the most dedicated universities, has been depleted to the extent that it falls far short of what is in fact needed for graduate study and research or for real-world application. This is true not only in the areas of theoretical work but also in the areas of computational matrix theory, which are becoming more and more important to the working researcher as personal computers become a common and powerful tool. Students are not only less able to formulate or even follow mathematical proofs, they are also less able to understand the underlying mathematics of the numerical algorithms they must use. The resulting knowledge gap has led to frustration and recrimination on the part of both students and faculty alike, with each silently – and sometimes not so silently – blaming the other for the resulting state of affairs. This book is written with the intention of bridging that gap. It was designed be used in one or more of several possible ways:

(1) As a self-study guide;

(2) As a textbook for a course in advanced linear algebra, either at the upper-class undergraduate level or at the first-year graduate level; or

(3) As a reference book.

It is also designed to be used to prepare for the linear algebra portion of prelim exams or PhD qualifying exams.

This volume is self-contained to the extent that it does not assume any previous knowledge of formal linear algebra, though the reader is assumed to have been exposed, at least informally, to some basic ideas or techniques, such as matrix manipulation and the solution of a small system of linear equations. It does, however, assume a seriousness of purpose, considerable

[1] This law, attributed to John Crow of King's College, London, is quoted by R. V. Jones in his book *Most Secret War*.

motivation, and modicum of mathematical sophistication on the part of the reader.

The book also contains a large number of exercises, many of which are quite challenging, which I have come across or thought up in over thirty years of teaching. Many of these exercises have appeared in print before, in such journals as *American Mathematical Monthly*, *College Mathematics Journal*, *Mathematical Gazette*, or *Mathematics Magazine*, in various mathematics competitions or circulated problem collections, or even on the internet. Some were donated to me by colleagues and even students, and some originated in files of old exams at various universities which I have visited in the course of my career. Since, over the years, I did not keep track of their sources, all I can do is offer a collective acknowledgement to all those to whom it is due. Good problem formulators, like the God of the abbot of Citeaux, know their own. Deliberately, difficult exercises are not marked with an asterisk or other symbol. Solving exercises is an integral part of learning mathematics and the reader is definitely expected to do so, especially when the book is used for self-study.

Solving a problem using theoretical mathematics is often very different from solving it computationally, and so strong emphasis is placed on the interplay of theoretical and computational results. Real-life implementation of theoretical results is perpetually plagued by errors: errors in modelling, errors in data acquisition and recording, and errors in the computational process itself due to roundoff and truncation. There are further constraints imposed by limitations in time and memory available for computation. Thus the most elegant theoretical solution to a problem may not lead to the most efficient or useful method of solution in practice. While no reference is made to particular computer software, the concurrent use of a personal computer equipped symbolic-manipulation software such as MAPLE, MATHEMATICA, MATLAB or MUPAD is definitely advised.

In order to show the "human face" of mathematics, the book also includes a large number of thumbnail photographs of researchers who have contributed to the development of the material presented in this volume.

Acknowledgements. Most of the first edition this book was written while the I was a visitor at the University of Iowa in Iowa City and at the University of California in Berkeley. I would like to thank both institutions for providing the facilities and, more importantly, the mathematical atmosphere which allowed me to concentrate on writing. This edition was extensively revised after I retired from teaching at the University of Haifa in April, 2004.

I have talked to many students and faculty members about my plans for this book and have obtained valuable insights from them. In particular, I would like to acknowledge the aid of the following colleagues and students who were kind enough to read the preliminary versions of this book and

offer their comments and corrections: Prof. Daniel Anderson (University of Iowa), Prof. Adi Ben-Israel (Rutgers University), Prof. Robert Cacioppo (Truman State University), Prof. Joseph Felsenstein (University of Washington), Prof. Ryan Skip Garibaldi (Emory University), Mr. George Kirkup (University of California, Berkeley), Prof. Earl Taft (Rutgers University), Mr. Gil Varnik (University of Haifa).

Photo credits. The photograph of Dr. Shmuel Winograd is used with the kind permission of the Department of Computer Science of the City University of Hong Kong. The photographs of Prof. Ben-Israel, Prof. Blass, Prof. Kublanovskaya, and Prof. Strassen are used with their respective kind permissions. The photograph of Prof. Greville is used with the kind permission of Mrs. Greville. The photograph of Prof. Rutishauser is used with the kind permission of Prof. Walter Gander. The photograph of Prof. V. N. Faddeeva is used with the kind permission of Dr. Vera Simonova. The photograph of Prof. Zorn is used with the kind permission of his son, Jens Zorn. The photograph of J. W. Givens was taken from a group photograph of the participants at the 1964 Gatlinburg Conference on Numerical Algebra. All other photographs are taken from the MacTutor History of Mathematics Archive website (http://www-history.mcs.st-andrews.ac.uk/history/index.html), the portrait gallery of mathematicians at the Trucsmatheux website (http://trucsmaths.free.fr/), or similar websites. To the best knowledge of the managers of those sites, and to the best of my knowledge, they are in the public domain.

Contents

1
Notation and terminology

Sets will be denoted by braces, $\{\ \ \}$, between which we will either enumerate the elements of the set or give a rule for determining whether something is an element of the set or not, as in $\{x \mid p(x)\}$, which is read "the set of all x such that $p(x)$". If a is an element of a set A we write $a \in A$; if it is not an element of A, we write $a \notin A$. When one enumerates the elements of a set, the order is not important. Thus $\{1,2,3,4\}$ and $\{4,1,3,2\}$ both denote the same set. However, we often do wish to impose an order on sets the elements of which we enumerate. Rather than introduce new and cumbersome notation to handle this, we will make the convention that when we enumerate the elements of a finite or countably-infinite set, we will assume an implied order, reading from left to right. Thus, the implied order on the set $\{1,2,3,\ldots\}$ is indeed the usual one. The empty set, namely the set having no elements, is denoted by $\varnothing$. Sometimes we will use the word "collection" as a synonym for "set", generally to avoid talking about "sets of sets".

A finite or countably-infinite selection of elements of a set A is a **list.** Members of a list are assumed to be in a definite order, given by their indices or by the implied order of reading from left to right. Lists are usually written without brackets: $a_1,\ldots,a_n$, though, in certain contexts, it will be more convenient to write them as ordered n-tuples $(a_1,\ldots,a_n)$. Note that the elements of a list need not be distinct: $3,1,4,1,5,9$ is a list of six positive integers, the second and fourth elements of which are equal to 1. A countably-infinite list of elements of a set A is also often called a **sequence** of elements of A. The set of all distinct members of a list is called the **underlying subset** of the list.

If A and B are sets, then their **union** $A \cup B$ is the set of all elements that belong to either A or B, and their **intersection** $A \cap B$ is the set of all elements belonging both to A and to B. More generally, if $\{A_i \mid i \in \Omega\}$ is a (possibly-infinite) collection of sets, then $\bigcup_{i\in\Omega} A_i$ is the set of all elements that belong to at least one of the A_i and $\bigcap_{i\in\Omega} A_i$ is the set of all elements that belong to all of the A_i. If A and B are sets, then the **difference set** $A \setminus B$ is the set of all elements of A which do not belong to B.

A **function** f from a nonempty set A to a nonempty set B is a rule which assigns to each element a of A a unique element $f(a)$ of B. The set A is called the **domain** of the function and the set B is called the **range** of the function. To denote that f is a function from A to B, we write $f : A \to B$. To denote that an element b of B is assigned to an element a of A by f, we write $f : a \mapsto b$. (Note the different form of the arrow!) This notation is particularly helpful in the case that the function f is defined by a formula. Thus, for example, if f is a function from the set of integers to the set of integers defined by $f : a \mapsto a^3$, then we know that f assigns to each integer its cube. The set of all functions from a nonempty set A to a nonempty set B is denoted by B^A. If $f \in B^A$ and if A' is a nonempty subset of A, then the **restriction** of f to A' is the function $f' : A' \to B$ defined by $f' : a' \mapsto f(a')$ for all $a' \in A'$.

Functions f and g in B^A are **equal** if and only if $f(a) = g(a)$ for all $a \in A$. In this case we write $f = g$. A function $f \in B^A$ is **monic** if and only if it assigns different elements of B to different elements of A, i.e. if and only if $f(a_1) \neq f(a_2)$ whenever $a_1 \neq a_2$ in A. A function $f \in B^A$ is **epic** if and only if every element of B is assigned by f to some element of A. A function which is both monic and epic is **bijective**. A bijective function from a set A to a set B determines a bijective correspondence between the elements of A and the elements of B. If $f : A \to B$ is a bijective function, then we can define the **inverse function** $f^{-1} : B \to A$ defined by the condition that $f^{-1}(b) = a$ if and only if $f(a) = b$. This inverse function is also bijective. A bijective function from a set A to itself is a **permutation** of A. Note that there is always at least one permutation of any nonempty set A, namely the identity function $a \mapsto a$.

The **cartesian product** $A_1 \times A_2$ of nonempty sets A_1 and A_2 is the set of all ordered pairs (a_1, a_2), where $a_1 \in A_1$ and $a_2 \in A_2$. More generally, if $A_1, \ldots, A_n$ is a list of nonempty sets, then $A_1 \times \ldots \times A_n$ is the set of all ordered n-tuples $(a_1, \ldots, a_n)$ satisfying the condition that $a_i \in A_i$ for each $1 \leq i \leq n$. Note that each ordered n-tuple $(a_1, \ldots, a_n)$

uniquely defines a function $f : \{1, \ldots, n\} \to \cup_{i=1}^{n} A_i$ given by $f : i \mapsto a_i$ for each $1 \leq i \leq n$. Conversely, each function $f : \{1, \ldots, n\} \to \bigcup_{i=1}^{n} A_i$ satisfying the condition that $f(i) \in A_i$ for $1 \leq i \leq n$, defines such an ordered n-tuple, namely $(f(1), \ldots, f(n))$. This suggests a method for defining the cartesian product of an arbitrary collection of nonempty sets. If $\{A_i \mid i \in \Omega\}$ is an arbitrary collection of nonempty sets, then the set $\prod_{i \in \Omega} A_i$ is defined to be the set of all those functions f from Ω to $\bigcup_{i \in \Omega} A_i$ satisfying the condition that $f(i) \in A_i$ for each $i \in \Omega$. The existence of such functions is guaranteed by a fundamental axiom of set theory, known as the **Axiom of Choice**. A certain amount of controversy surrounds this axiom, and there are mathematicians who prefer to make as little use of it as possible. However, we will need it constantly throughout this book, and so will always assume that it holds.

In the foregoing construction we did not assume that the sets A_i were necessarily distinct. Indeed, it may very well happen that there exists a set A such that $A_i = A$ for all $i \in \Omega$. In that case, we see that $\prod_{i \in \Omega} A_i$ is just A^{Ω}. If the set Ω is finite, say $\Omega = \{1, \ldots, n\}$, then we write A^n instead of A^{Ω}. Thus, A^n is just the set of all ordered n-tuples $(a_1, \ldots, a_n)$ of elements of A.

We use the following standard notation for some common sets of numbers

$\mathbb{N}$	the set of all nonnegative integers
$\mathbb{Z}$	the set of all integers
$\mathbb{Q}$	the set of all rational numbers
$\mathbb{R}$	the set of all real numbers
$\mathbb{C}$	the set of all complex numbers

Other notion is introduced throughout the text, as is appropriate. See the Summary of Notation at the end of the book.

2
Fields

The way of mathematical thought is twofold: the mathematician first proceeds inductively from the particular to the general and then deductively from the general to the particular. Moreover, throughout its development, mathematics has shown two aspects – the conceptual and the computational – the symphonic interleaving of which forms one of the major aspects of the subject's aesthetic.

Let us therefore begin with the first mathematical structure: numbers. By the Hellenistic times, mathematicians distinguished between two types of numbers: the **rational** numbers, namely those which could be written in the form $\frac{m}{n}$ for some integer m and some nonnegative integer n, and those numbers representing the geometric magnitude of segments of the line, which today we call **real** numbers and which, in decimal notation, are written in the form $m.k_1k_2k_3\ldots$ where m is an integer and the k_i are digits. The fact that the set $\mathbb{Q}$ of rational numbers is not equal to the set $\mathbb{R}$ of real numbers was already noticed by the followers of the mathematician/mystic Pythagoras. On both sets of numbers we define operations of addition and multiplication which satisfy certain rules of manipulation. Isolating these rules as part of a formal system was a task first taken on in earnest by nineteenth-century British and German mathematicians. From their studies evolved the notion of a field, which will be basic to our considerations. However, since fields are not our primary object of study, we will delve only minimally into this fascinating notion. A serious consideration of field theory must be deferred to an advanced course in abstract algebra.[1]

A nonempty set F together with two functions $F \times F \to F$, respectively called **addition** (as usual, denoted by $+$) and **multiplication** (as usual, denoted by $\cdot$ or by concatenation), is a **field** if the following conditions are satisfied:

(1) (**associativity of addition and multiplication**): $a + (b + c) = (a + b) + c$ and $a(bc) = (ab)c$ for all $a, b, c \in F$.

(2) (**commutativity of addition and multiplication**): $a + b = b + a$ and $ab = ba$ for all $a, b \in F$.

(3) (**distributivity of multiplication over addition**): $a(b + c) = ab + ac$ for all $a, b, c \in F$.

(4) (**existence of identity elements for addition and multiplication**): There exist distinct elements of F, which we will denote by 0 and 1 respectively, satisfying $a + 0 = a$ and $a1 = a$ for all $a \in F$.

(5) (**existence of additive inverses**): For each $a \in F$ there exists an element of F, which we will denote by $-a$, satisfying $a + (-a) = 0$.

(6) (**existence of multiplicative inverses**): For each $0 \neq a \in F$ there exists an element of F, which we will denote by a^{-1}, satisfying $a^{-1}a = 1$.

Note that we did not assume that the elements $-a$ and a^{-1} are unique, though we will soon prove that in fact they are. If a and b are elements of a field F, we will follow the usual conventions by writing $a - b$ instead of $a + (-b)$ and $\frac{a}{b}$ instead of ab^{-1}. Moreover, if $0 \neq a \in F$ and if n is a positive integer, then na denotes the sum $a + \ldots + a$ (n summands) and a^n denotes the product $a \cdot \ldots \cdot a$ ($n \cdot$ factors). If n is a negative integer, then na denotes $(-n)(-a)$ and a^n denotes $(a^{-1})^{-n}$. Finally, if $n = 0$ then na denotes the field element 0 and a^n denotes the field element 1. For $0 = a \in F$, we define $na = 0$ for all integers n and

---

[1] 

The development of the abstract theory of fields is generally credited to the 19th-century German mathematician **Heinrich Weber**, based on earlier work by the German mathematicians **Richard Dedekind** and **Leopold Kronecker**. Another 19th-century mathematician, the British **Augustus De Morgan**, was the first to isolate the importance of such properties as associativity, distributivity, and so forth. The final axioms of a field are due to the 20th-century German mathematician **Ernst Steinitz**.

$a^n = 0$ for all positive integers n. The symbol 0^k is not defined for $k \leq 0$.

As an immediate consequence of the associativity and commutativity of addition, we see that the sum of any list $a_1, \ldots, a_n$ of elements of a field F is the same, no matter in which order we add them. We can therefore unambiguously write $a_1 + \ldots + a_n$. This sum is also often denoted by $\sum_{i=1}^{n} a_i$. Similarly, the product of these elements is the same, no matter in which order we multiply them. We can therefore unambiguously write $a_1 \cdot \ldots \cdot a_n$. This product is also often denoted by $\prod_{i=1}^{n} a_i$. Also, a simple inductive argument shows that multiplication distributes over arbitrary sums: if $a \in F$ and $b_1, \ldots b_n$ is a list of elements of F then $a\left(\sum_{i=1}^{n} b_i\right) = \sum_{i=1}^{n} ab_i$.

We easily see that $\mathbb{Q}$ and $\mathbb{R}$, with the usual addition and multiplication, are fields.

A subset G of a field F is a **subfield** if and only if it contains 0 and 1, is closed under addition and multiplication, and contains the additive and multiplicative inverses of all of its nonzero elements. Thus, for example, $\mathbb{Q}$ is a subfield of $\mathbb{R}$. The intersection of a collection of subfields of a field F is again a subfield of F.

We now want to look at several additional important examples of fields.

Example: Let $\mathbb{C} = \mathbb{R}^2$ and define operations of addition and multiplication on $\mathbb{C}$ by setting $(a,b)+(c,d) = (a+c, b+d)$ and $(a,b)\cdot(c,d) = (ac-bd, ad+bc)$. These operations define the structure of a field on $\mathbb{C}$, in which the identity element for addition is $(0,0)$, the identity element for multiplication is $(1,0)$, the additive inverse of (a,b) is $(-a,-b)$, and

$$(a,b)^{-1} = \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$$

for all $(0,0) \neq (a,b)$. This field is called the field of **complex numbers**. The set of all elements of $\mathbb{C}$ of the form $(a,0)$ forms a subfield of $\mathbb{C}$, which we normally identify with $\mathbb{R}$ and therefore it is standard to consider $\mathbb{R}$ as a subfield of $\mathbb{C}$. In particular, we write a instead of $(a,0)$ for any real number a. The element $(0,1)$ of $\mathbb{C}$ is denoted by i. This element satisfies the condition that $i^2 = (-1,0)$ and so it is often written as $\sqrt{-1}$. We also note that any element (a,b) of $\mathbb{C}$ can be written as $(a,0) + b(0,1) = a + bi$, and, indeed, that is the way complex numbers are usually written and how we will denote them from now on. If $z = a + bi$, then a is the **real part** of z, which is often denoted by $\mathrm{Re}(z)$, while bi is the **imaginary part** of z, which is often denoted by $\mathrm{Im}(z)$. The