LINEAR ALGEBRA FOR EVERYONE

GILBERT STRANG
Massachusetts Institute of Technology

WELLESLEY - CAMBRIDGE PRESS
Box 812060 Wellesley MA 02482

Linear Algebra for Everyone
Copyright ©2020 by Gilbert Strang
ISBN 978-1-7331466-3-0

LaTeX **typesetting by Ashley C. Fernandes** (info@problemsolvingpathway.com)
Printed in the United States of America

9 8 7 6 5 4 3 2 1

QA184.2 .S773 2020 | DDC 512/.5–dc23

Other texts from Wellesley - Cambridge Press

Linear Algebra and Learning from Data, 2019, Gilbert Strang ISBN 978-0-6921963-8-0

Introduction to Linear Algebra, 5th Ed., 2016, Gilbert Strang ISBN 978-0-9802327-7-6

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, 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
LAFEeveryone@gmail.com

Gilbert Strang's page : **math.mit.edu/∼gs**
For orders : **math.mit.edu/weborder.php**
Outside US/Canada : **www.cambridge.org**
Select books, India : **www.wellesleypublishers.com**

The website for this book (with Solution Manual) is **math.mit.edu/everyone**
2019 book : **Linear Algebra and Learning from Data** (**math.mit.edu/learningfromdata**)
2016 book : **Introduction to Linear Algebra, 5th Edition** (**math.mit.edu/linearalgebra**)
2014 book : **Differential Equations and Linear Algebra** (**math.mit.edu/dela**)

Linear Algebra is included in MIT's OpenCourseWare **ocw.mit.edu/courses/mathematics**
Those videos (including 18.06SC and 18.065) are also on **www.youtube.com/mitocw**
18.06 Linear Algebra 18.06SC with problem solutions 18.065 Learning from Data

The cover design was created by Gail Corbett and Lois Sellers : **lsellersdesign.com**

1.4 Matrix Multiplication and $A = CR$

1 To multiply AB we need *row length for* $A =$ *column length for* B.

2 The number in row i, column j of AB is (**row i of A**) $\cdot$ (**column j of B**).

3 By columns: **A times column j of B produces column j of AB.**

4 Usually AB is different from BA. But always $(\boldsymbol{AB})\,\boldsymbol{C} = \boldsymbol{A}\,(\boldsymbol{BC})$.

5 If A has r independent columns, then $\boldsymbol{A} = \boldsymbol{CR} = (m \text{ by } r)\,(r \text{ by } n)$.

At this point we can multiply a matrix A times a vector $\boldsymbol{x}$ to produce $A\boldsymbol{x}$. Remember the row way and the column way. The output is a vector.

Row way Dot products of $\boldsymbol{x}$ with each row of A

Column way $A\boldsymbol{x} = x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n =$ combination of the columns of A

Now we come to the higher level operation in linear algebra: **Multiply two matrices**. We can multiply AB if their shapes are right. When A has n columns, B must have n rows.

If A is m by n and B is n by p, then AB is m by p: m columns and p rows.

The rules for AB will be an extension of the rules for $A\boldsymbol{x}$. We can think of the vector $\boldsymbol{x}$ as a matrix B with only one column. What we did for $A\boldsymbol{x}$ we will now do for AB.

The columns of B are vectors $\begin{bmatrix} \boldsymbol{x} & \boldsymbol{y} & \boldsymbol{z} \end{bmatrix}$. **The columns of AB are vectors** $\begin{bmatrix} A\boldsymbol{x} & A\boldsymbol{y} & A\boldsymbol{z} \end{bmatrix}$

In other words, **multiply A times each column of B**. There are two ways to multiply A times a column vector, and those give *two ways to multiply AB*:

Dot products	(**row i of A**) $\cdot$ (**column j of B**) goes into row i, column j of AB
Combinations of columns of A	Use the numbers from each column of B

We have dot products (numbers) or linear combinations of columns of A (vectors). For computing by hand, I would use the row way to find each number in AB. I "think" the column way to see the big picture: **Columns of AB are combinations of columns of A**.

Example 1 Multiply $AB = \begin{bmatrix} \mathbf{1} & \mathbf{2} \\ \mathbf{3} & \mathbf{4} \end{bmatrix} \begin{bmatrix} \mathbf{5} & \mathbf{6} \\ \mathbf{7} & \mathbf{8} \end{bmatrix}$ both ways. How many steps?

The dot product (row 1 of A) $\cdot$ (column 1 of B) is $(\mathbf{1}, \mathbf{2}) \cdot (\mathbf{5}, \mathbf{7}) = 5 + 14 = \mathbf{19}$

(**Rows of A**) $\cdot$ (**columns of B**) $\quad AB = \begin{bmatrix} \text{row 1} \cdot \text{col 1} & \text{row 1} \cdot \text{col 2} \\ \text{row 2} \cdot \text{col 1} & \text{row 2} \cdot \text{col 2} \end{bmatrix} = \begin{bmatrix} \mathbf{19} & \mathbf{22} \\ \mathbf{43} & \mathbf{50} \end{bmatrix}$

$A\boldsymbol{b}_1$ and $A\boldsymbol{b}_2$ are combinations of the columns of A $\quad AB = \left[\mathbf{5}\begin{bmatrix} 1 \\ 3 \end{bmatrix} + \mathbf{7}\begin{bmatrix} 2 \\ 4 \end{bmatrix} \quad \mathbf{6}\begin{bmatrix} 1 \\ 3 \end{bmatrix} + \mathbf{8}\begin{bmatrix} 2 \\ 4 \end{bmatrix} \right] = \begin{bmatrix} \mathbf{19} & \mathbf{22} \\ \mathbf{43} & \mathbf{50} \end{bmatrix}$

Either way, AB requires 8 multiplications for 2 by 2 matrices. Multiplying $A\boldsymbol{x}$ needed mn multiplications for an m by n matrix. When B has p columns, we have to multiply A times each column. Then *multiplying AB uses mnp multiplications* for (m by n) times (n by p). When $m = n = p = 2$, we have $2 \cdot 2 \cdot 2 = 8$ multiplications.

A difficult question: Could n by n matrices A and B be multiplied with fewer than n^3 small multiplications? *This is known to be possible* (allowing extra additions). Right now we don't know the smallest exponent E in the multiplication count n^E. We know that $E < 3$ and in fact $E < 2.373$. But $E = 2.0001$ may be impossible.

Here is a challenge question about matrix multiplication. *Explain why every vector in* $\mathbf{C}(AB)$—every combination of the columns of AB—*is also in the column space of A.*

Example 2 The **identity matrix** I has $AI = A$ and $IB = B$ if matrix sizes are right.

$$\boldsymbol{AI} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \boldsymbol{A} \qquad \text{for every } A$$

$$\boldsymbol{IB} = \begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{1} \end{bmatrix} \begin{bmatrix} P & Q \\ R & S \end{bmatrix} = \begin{bmatrix} P & Q \\ R & S \end{bmatrix} = \boldsymbol{B} \qquad \text{for every } B$$

Example 3 The matrix $\boldsymbol{E} = \begin{bmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{bmatrix}$ will *exchange columns* or *exchange rows*.

$$\boldsymbol{AE} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ d & c \end{bmatrix} \qquad \text{Exchange columns of } A \ (E \text{ is on the right})$$

$$\boldsymbol{EB} = \begin{bmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{bmatrix} \begin{bmatrix} P & Q \\ R & S \end{bmatrix} = \begin{bmatrix} R & S \\ P & Q \end{bmatrix} \qquad \text{Exchange rows of } B \ (E \text{ is on the left})$$

Example 4 $AE \neq EA$ for most matrices: Exchange columns or exchange rows.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad AE = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix} \text{ is not the same as } EA = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$$

This example shows an important fact: $\boldsymbol{AB}$ **can easily be different from** $\boldsymbol{BA}$. Matrix multiplication is not "*commutative*". We must keep matrices AB or ABC in order. There are certainly examples where $AB = BA$, but those special cases are not typical.

Example 5 Squaring the exchange matrix gives $E^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = I$. Why is this true?

Example 6 $\boldsymbol{AB}$ **times** $\boldsymbol{C}$ **equals** $\boldsymbol{A}$ **times** $\boldsymbol{BC}$. Matrix multiplication is "*associative*".

I include this here without proof, because it is so important: $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$. The order ABC of those matrices must stay the same! But we can multiply AB first or multiply BC first. Many many proofs in linear algebra depend on this simple fact.

Certainly the matrix sizes for A, B, C must match: (m by n) $\times$ (n by p) $\times$ (p by q). A special 2 by 2 case of this associative law would be $(EA)E = E(AE)$. Exchange the rows of A first, or exchange the columns of A first. The triple product EAE does both.

Rank One Matrices and $\boldsymbol{A = CR}$

All columns of a rank one matrix lie on the same line. That line is the column space $\mathbf{C}(A)$. Examples in Section 1.3 pointed to a remarkable fact: *The rows also lie on a line*. When all the columns are in the same column direction, then all the rows are in the same row direction. Here is an example:

$$A = \begin{bmatrix} 1 & 2 & 10 & 100 \\ 3 & 6 & 30 & 300 \\ 2 & 4 & 20 & 200 \end{bmatrix} = \begin{array}{l} \text{rank one matrix} \\ \text{one independent column} \\ \text{one independent row !} \end{array}$$

All columns are multiples of $(1,3,2)$. All rows are multiples of $\begin{bmatrix} 1 & 2 & 10 & 100 \end{bmatrix}$. Only one independent row when there is only one independent column. *Why is this true* ?

Our approach is through matrix multiplication. We factor A into C times R. For this special matrix, C has one column and R has one row. $\boldsymbol{CR}$ is $(3 \times 1)\,(1 \times 4)$.

$$\boxed{\boldsymbol{A = CR} \qquad \begin{bmatrix} 1 & 2 & 10 & 100 \\ 3 & 6 & 30 & 300 \\ 2 & 4 & 20 & 200 \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{3} \\ \mathbf{2} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{2} & \mathbf{10} & \mathbf{100} \end{bmatrix}} \qquad (1)$$

The dot products (row of C)$\cdot$(column of R) are just multiplications like 3 times 10. This is multiplication of thin matrices CR. Only 12 small multiplications.

The rows of A are numbers $1,3,2$ times the (only) row $\begin{bmatrix} 1 & 2 & 10 & 100 \end{bmatrix}$ of R. By factoring this special A into **one column times one row**, the conclusion jumps out:

If the column space of $\boldsymbol{A}$ is a line, the row space of $\boldsymbol{A}$ is also a line.

One column in C, one row in R. That is beautiful, but we are certainly not finished. Our big goal is to allow r columns in C and to find r rows in R. And to see $A = CR$.

$\boldsymbol{C}$ Contains Independent Columns

Suppose we go from left to right, looking for independent columns in any matrix A:

> If column 1 of A is not all zero, put it into the matrix C
>
> If column 2 of A is not a multiple of column 1, put it into C
>
> If column 3 of A is not a combination of columns 1 and 2, put it into C. *Continue*.

At the end C will have r columns taken from A. That number r is the **rank of $\boldsymbol{A}$**. The n columns of A might be dependent. The r columns of C will surely be **independent**.

Independent	*No column of C is a combination of previous columns*
columns	*No combination of columns gives* $C\boldsymbol{x} = \mathbf{0}$ *except* $\boldsymbol{x} =$ *all zeros*

When those independent columns combine to give all columns, we have a **basis**.

$C\boldsymbol{x} = \mathbf{0}$ means that x_1(column 1 of C) $+$ x_2(column 2 of C) $+ \cdots =$ *zero vector*. With independent columns, this only happens if *all x's are zero*. Otherwise we can divide by the last nonzero coefficient x and that column would be a combination of the earlier columns—which our construction forbids. Therefore C has independent columns.

$$\textbf{Example 7} \qquad A = \begin{bmatrix} \mathbf{2} & 6 & \mathbf{4} \\ \mathbf{4} & 12 & \mathbf{8} \\ \mathbf{1} & 3 & \mathbf{5} \end{bmatrix} \quad \text{leads to} \quad C = \begin{bmatrix} \mathbf{2} & \mathbf{4} \\ \mathbf{4} & \mathbf{8} \\ \mathbf{1} & \mathbf{5} \end{bmatrix}$$

Column 1 goes into C. Column 2 does not (3 times column 1). Column 3 goes into C.

Matrix Multiplication C times R

Now comes the new step to $\boldsymbol{A = CR}$. $\boldsymbol{R}$ tells how to produce the columns of $\boldsymbol{A}$ from the columns of $\boldsymbol{C}$. The first column of A is actually in C, so the first column of R just has 1 and 0. The third column of A is also in C, so the third column of R just has 0 and 1.

$$\begin{matrix}\textbf{Rank 2}\\ \textbf{Notice } \boldsymbol{I}\\ \textbf{inside } \boldsymbol{R}\end{matrix} \qquad A = \begin{bmatrix} 2 & 6 & 4 \\ 4 & 12 & 8 \\ 1 & 3 & 5 \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{4} \\ \mathbf{4} & \mathbf{8} \\ \mathbf{1} & \mathbf{5} \end{bmatrix} \begin{bmatrix} \mathbf{1} & ? & \mathbf{0} \\ \mathbf{0} & ? & \mathbf{1} \end{bmatrix} = \boldsymbol{CR}. \qquad (2)$$

Two columns of A went straight into C, so *part of R is the identity matrix*. The question marks are in column 2 because column 2 of A is *not in C*. It was not an independent column. Column 2 of A is 3 times column 1. *That number* 3 *goes into R*. Then R shows how to combine the two columns of C to get all three columns of the original A.

$$\begin{matrix} A \text{ is } m \times n \\ C \text{ is } m \times r \\ R \text{ is } r \times n \end{matrix} \qquad \boxed{\boldsymbol{A = CR} \text{ is } \begin{bmatrix} 2 & 6 & 4 \\ 4 & 12 & 8 \\ 1 & 3 & 5 \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{4} \\ \mathbf{4} & \mathbf{8} \\ \mathbf{1} & \mathbf{5} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{3} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}} \qquad (3)$$

This completes $A = CR$. The magic is now seen in the rows. **All the rows of $\boldsymbol{A}$ come from the rows of $\boldsymbol{R}$.** This fact follows immediately from matrix multiplication CR:

$$\begin{matrix}\textbf{Multiply } \boldsymbol{CR}\\ \textbf{using rows of } \boldsymbol{R}\end{matrix} \qquad \begin{matrix} \text{Row 1 of } A \text{ is} & \mathbf{2}\,(\text{row 1 of } R) + \mathbf{4}\,(\text{row 2 of } R) \\ \text{Row 2 of } A \text{ is} & \mathbf{4}\,(\text{row 1 of } R) + \mathbf{8}\,(\text{row 2 of } R) \\ \text{Row 3 of } A \text{ is} & \mathbf{1}\,(\text{row 1 of } R) + \mathbf{5}\,(\text{row 2 of } R) \end{matrix}$$

$\boldsymbol{R}$ has the same row space as $\boldsymbol{A}$. Combinations of rows of R produce every row of A. This A has only 2 independent rows, not 3. Two rows in R combine to give all rows of A.

$$\begin{matrix}\textbf{Second example}\\ \textbf{of } \boldsymbol{A = CR}\\ \textbf{from the front cover}\end{matrix} \qquad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{-1} \\ \mathbf{0} & \mathbf{1} & \mathbf{2} \end{bmatrix} \qquad (4)$$

When a column of A goes into C, a column of I goes into R. The "free" column $-1, 2$ of R tells us how to produce the *dependent* column of A from the *independent* columns in C. Column 3 of A is $\mathbf{-1}\,(1,4,7) + \mathbf{2}\,(2,5,8)$. And also: rows of A from rows of R.

Column $\boldsymbol{j}$ of $\boldsymbol{A}$ $=$ **$\boldsymbol{C}$** times **column $\boldsymbol{j}$ of $\boldsymbol{R}$**. **Row $\boldsymbol{i}$ of $\boldsymbol{A}$** $=$ **row $\boldsymbol{i}$ of $\boldsymbol{C}$** times **$\boldsymbol{R}$**.

Question If all n columns of A are independent, then $C = A$. What matrix is R?

Answer This case of n independent columns has $\boldsymbol{R = I}$ (identity matrix). The rank is n.

How to find $\boldsymbol{R}$. Start with r independent columns of A going into C.

If column 3 of $A =$ 2nd independent column in C, then column 3 of R is $\begin{matrix}\mathbf{0}\\ \mathbf{1}\end{matrix}$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & \mathbf{3} & 4 \\ 1 & 2 & \mathbf{4} & 5 \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{3} \\ 1 & \mathbf{4} \end{bmatrix} \begin{bmatrix} 1 & 2 & \mathbf{0} & 1 \\ 0 & 0 & \mathbf{1} & 1 \end{bmatrix} = \boldsymbol{CR} \qquad \text{All three ranks} = 2$$

Dependent: If column 4 of $A =$ columns $1 + 2$ of C, then column 4 of R is $\begin{matrix}\mathbf{1}\\ \mathbf{1}\end{matrix}$

$\boldsymbol{R}$ tells how to recover all columns of $\boldsymbol{A}$ from the independent columns in $\boldsymbol{C}$.

Here is an informal proof that ***row rank of A equals column rank of A***

1. The r columns of C are independent (by their construction)
2. Every column of A is a combination of those r columns of C (because $A = CR$)
3. The r rows of R are independent (they contain the r by r matrix I)
4. Every row of A is a combination of those r rows of R (because $A = CR$)

Key facts

The r columns of C are a **basis** for the column space of A: **dimension $\boldsymbol{r}$**

The r rows of R are a **basis** for the row space of A: **dimension $\boldsymbol{r}$**

Those words "basis" and "dimension" are properly defined in Section 3.4! Section 3.2 will show how the same row matrix R can be constructed directly from the "reduced row echelon form" of A, by deleting any zero rows. Chapter 1 starts with independent columns of A, placed in C. *Chapter* 3 *starts with rows of A, and combines them into R.*

We are emphasizing CR because both matrices are important. C contains r independent columns of A. R tells how to combine those columns to give all columns of A. (R contains I, when columns of A are already in C.) Chapter 3 will produce R directly from A by *elimination*, the most used algorithm in computational mathematics. This will be the key to a fundamental problem: solving linear equations $A\boldsymbol{x} = \boldsymbol{b}$.

Why is Matrix Multiplication AB Defined This Way?

The definition of AB was chosen to produce this crucial equation: **$\boldsymbol{(AB)}$ times $\boldsymbol{x}$ is equal to $\boldsymbol{A}$ times $\boldsymbol{Bx}$.** This leads to the all-important law $\boldsymbol{(AB)C = A(BC)}$. We had no other reasonable choice for AB! Linear algebra will use these laws over and over. Let me show in three steps why that crucial equation $(AB)\boldsymbol{x} = A(B\boldsymbol{x})$ is correct:

$B\boldsymbol{x}$ is a combination $x_1\boldsymbol{b}_1 + x_2\boldsymbol{b}_2 + \cdots + x_n\boldsymbol{b}_n$ of the columns of B.

Matrix-vector multiplication is linear: $A(B\boldsymbol{x}) = x_1 A\boldsymbol{b}_1 + x_2 A\boldsymbol{b}_2 + \cdots + x_n(A\boldsymbol{b}_n)$.

We want this to agree with $(AB)\boldsymbol{x} = x_1(\text{column 1 of } AB) + \cdots + x_n(\text{column } n \text{ of } AB)$.

Compare lines 2 and 3. *Column* 1 *of AB absolutely must equal A times column* 1 *of B.* This is our rule: When $B = \begin{bmatrix} \boldsymbol{x} & \boldsymbol{y} & \boldsymbol{z} \end{bmatrix}$ the columns of AB are $\begin{bmatrix} A\boldsymbol{x} & A\boldsymbol{y} & A\boldsymbol{z} \end{bmatrix}$.

Example 8 $\quad A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \quad \boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (AB)\boldsymbol{x} = A(B\boldsymbol{x})$

When we show that $(AB)\boldsymbol{x} = A(B\boldsymbol{x})$, that fact will allow us to erase the parentheses. *We can just write* $AB\boldsymbol{x}$. For three matrices we can just write ABC.

To compare $(AB)\boldsymbol{x}$ with $A(B\boldsymbol{x})$, remember that Example 1 computed AB :

$$(AB)\boldsymbol{x} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{19} \\ \mathbf{43} \end{bmatrix}$$

$$A(B\boldsymbol{x}) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} = \begin{bmatrix} \mathbf{19} \\ \mathbf{43} \end{bmatrix}$$

The parentheses don't matter but the order ABC certainly does matter. The multiplications BAC and ACB almost always give different answers. In fact BAC may be impossible.

Columns of A times Rows of B

Before this chapter ends, I want to add this message. There is another way to multiply matrices (producing the same matrix AB as always). This way is not so well known, but it is powerful. **It multiplies columns of A times rows of B.** We will see it again and use it.

$$\boxed{AB = \underbrace{\begin{bmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{bmatrix}}_{\text{columns } \boldsymbol{a}_k} \underbrace{\begin{bmatrix} \boldsymbol{b}_1^* \\ \vdots \\ \boldsymbol{b}_n^* \end{bmatrix}}_{\text{rows } \boldsymbol{b}_k^*} = \underbrace{\boldsymbol{a}_1\boldsymbol{b}_1^* + \boldsymbol{a}_2\boldsymbol{b}_2^* + \cdots + \boldsymbol{a}_n\boldsymbol{b}_n^*}_{\text{columns } \boldsymbol{a}_k \text{ times rows } \boldsymbol{b}_k^*}.} \qquad (5)$$

Those matrices $\boldsymbol{a}_k\boldsymbol{b}_k^*$ are called *outer products*. We recognize that they have rank one: column times row. They are entirely different from dot products (rows times columns, also known as *inner products*). If A is an m by n matrix and B is an n by p matrix, adding columns times rows gives the same answer AB as rows times columns.

Actually they involve the same mnp small multiplications but in a different order!

(**Row**) · (**Column**)	mp dot products, n multiplications each	**total** $\boldsymbol{mnp}$
(**Column**) (**Row**)	n rank one matrices, mp multiplications each	**total** $\boldsymbol{mnp}$

Columns × Rows for A times B

$$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 7 & 8 & 9 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \begin{bmatrix} 10 & 11 & 12 \end{bmatrix}$$

Rank 1 +Rank 1

$$= \begin{bmatrix} 7 & 8 & 9 \\ 14 & 16 & 18 \\ 21 & 24 & 27 \end{bmatrix} + \begin{bmatrix} 40 & 44 & 48 \\ 50 & 55 & 60 \\ 60 & 66 & 72 \end{bmatrix} = \begin{bmatrix} \mathbf{47} & \mathbf{52} & \mathbf{57} \\ \mathbf{64} & \mathbf{71} & \mathbf{78} \\ \mathbf{81} & \mathbf{90} & \mathbf{99} \end{bmatrix} \qquad (6)$$

This example has $mnp = (3)\,(2)\,(3) = 18$. At the start of the second line you see the 18 multiplications (in two 3 by 3 matrices). Then 9 additions give the correct answer AB.

As we learned in this section, the rank of AB is 2. *Two independent columns, not three. Two independent rows, not three*. The next chapter will use different words. *AB has no inverse matrix*: *it is not invertible*. And later in the book: *The determinant of AB is zero*.

Note about the "echelon matrices" $\boldsymbol{R}$ and $\boldsymbol{R_0}$

We were amazed to learn that the row matrix R in $A = CR$ is already a famous matrix in linear algebra! It is essentially the **"reduced row echelon form"** of the original A. MATLAB calls it **rref** (A) and includes $m - r$ zero rows. With the zero rows, we call it $\boldsymbol{R_0}$.

The factorization $\boldsymbol{A = CR}$ is a big step in linear algebra. The Problem Set will look closely at the matrix R, its form is remarkable. R has the identity matrix in r columns. Then C multiplies each column of R to produce a column of A. $\boldsymbol{R_0}$ **comes in Chapter 3.**

Example 9 $$A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{3a}_1 + \boldsymbol{4a}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \end{bmatrix} \begin{bmatrix} \mathbf{1} & 0 & \mathbf{3} \\ 0 & \mathbf{1} & \mathbf{4} \end{bmatrix} = \boldsymbol{CR}.$$

Here $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are the independent columns of A. The third column is dependent—a combination of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. Therefore it is in the plane produced by columns 1 and 2. All three matrices A, C, R have rank $r = 2$.

We can try that new way (**columns** $\times$ **rows**) to quickly multiply CR in Example 9:

Columns of $\boldsymbol{C}$ times rows of $\boldsymbol{R}$ $$CR = \boldsymbol{a}_1 \begin{bmatrix} 1 & 0 & 3 \end{bmatrix} + \boldsymbol{a}_2 \begin{bmatrix} 0 & 1 & 4 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & 3\boldsymbol{a}_1 + 4\boldsymbol{a}_2 \end{bmatrix} = A$$

Four Ways to Multiply $\boldsymbol{AB = C}$

$$\begin{bmatrix} \boldsymbol{X} & \boldsymbol{X} \\ \boldsymbol{x} & \boldsymbol{x} \\ \boldsymbol{x} & \boldsymbol{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} & \boldsymbol{x} & \boldsymbol{x} \\ \boldsymbol{X} & \boldsymbol{x} & \boldsymbol{x} & \boldsymbol{x} \end{bmatrix}$$ (**Row $\boldsymbol{i}$** of A) $\cdot$ (**Column $\boldsymbol{k}$** of B) $=$ **Number $\boldsymbol{C_{ik}}$**
$i = 1$ to 3 $\quad k = 1$ to 4 $\quad$ **12 numbers**

$$\begin{bmatrix} \boldsymbol{X} & \boldsymbol{X} \\ \boldsymbol{X} & \boldsymbol{X} \\ \boldsymbol{X} & \boldsymbol{X} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} & \boldsymbol{x} & \boldsymbol{x} \\ \boldsymbol{X} & \boldsymbol{x} & \boldsymbol{x} & \boldsymbol{x} \end{bmatrix}$$ $\boldsymbol{A}$ times (**Column $\boldsymbol{k}$** of B) $=$ **Column $\boldsymbol{k}$ of $\boldsymbol{C}$**
$k = 1$ to 4 $\quad$ **4 columns**

$$\begin{bmatrix} \boldsymbol{X} & \boldsymbol{X} \\ \boldsymbol{x} & \boldsymbol{x} \\ \boldsymbol{x} & \boldsymbol{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \\ \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \end{bmatrix}$$ (**Row $\boldsymbol{i}$** of A) times $\boldsymbol{B}$ $=$ **Row $\boldsymbol{i}$ of $\boldsymbol{C}$**
$i = 1$ to 3 $\quad$ **3 rows**

$$\begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{X} & \boldsymbol{x} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \\ \boldsymbol{x} & \boldsymbol{x} & \boldsymbol{x} & \boldsymbol{x} \end{bmatrix}$$ (**Column $\boldsymbol{j}$** of A) (**Row $\boldsymbol{j}$** of B) $=$ **Rank 1 Matrix**
$j = 1$ to 2 $\quad$ **2 matrices**

Problem Set 1.4

1 Construct this four-way table when A is m by n and B is n by p. How many dot products and columns and rows and rank one matrices go into AB? In all four cases the total count of small multiplications is mnp.

2 If all columns of $A = \begin{bmatrix} \boldsymbol{a} & \boldsymbol{a} & \boldsymbol{a} \end{bmatrix}$ contain the same $\boldsymbol{a} \neq \mathbf{0}$, what are C and R?

3 Multiply A times B (3 examples) using *dot products*: each row times each column.

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \qquad \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

4 Test the truth of the associative law $(AB)C = A(BC)$.

(a) $\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$

5 Why is it impossible for a matrix A with 7 columns and 4 rows to have 5 independent columns ? This is not a trivial or useless question.

6 Going from left to right, put each column of A into the matrix C if that column is not a combination of earlier columns :

$$A = \begin{bmatrix} 2 & -2 & 1 & 6 & 0 \\ 1 & -1 & 0 & 2 & 0 \\ 3 & -3 & 0 & 6 & 1 \end{bmatrix} \qquad C = \begin{bmatrix} 2 & & & \\ 1 & & & \\ 3 & & & \end{bmatrix}$$

7 Find R in Problem 6 so that $A = CR$. If your C has r columns, then R has r rows. The 5 columns of R tell how to produce the 5 columns of A from the columns in C.

8 This matrix A has 3 independent columns. So C has the same 3 columns as A. What is the 3 by 3 matrix R so that $A = CR$? What is different about B ?

$$\text{Upper triangular} \qquad A = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 4 & 4 \\ 0 & 0 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 0 & 4 \\ 0 & 0 & 6 \end{bmatrix}$$

9 Suppose A is a random 4 by 4 matrix. The probability is 1 that the columns of A are "independent". In that case, what are the matrices C and R in $A = CR$?

Note Random matrix theory has become an important part of applied linear algebra—especially for very large matrices when even multiplication AB is too expensive. An example of "*probability* 1" is choosing two whole numbers at random. The probability is 1 that they are different. But they could be the same ! Problem 10 is another example of this type.

10 Suppose A is a random 4 by 5 matrix. With probability 1, what can you say about C and R in $A = CR$? In particular, which columns of A (going into C) are probably independent of previous columns, going from left to right ?

11 Create your own example of a 4 by 4 matrix A of rank $r = 2$. Then factor A into $CR = (4 \text{ by } 2)\,(2 \text{ by } 4)$.

12 Factor these matrices into $A = CR = (m \text{ by } r)\,(r \text{ by } n)$: all ranks equal to r.

$$A_1 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 4 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 0 & 1 & 3 & 5 \end{bmatrix} \quad A_3 = \begin{bmatrix} 2 & 1 & 3 \\ 6 & 3 & 9 \end{bmatrix} \quad A_4 = \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 2 & 2 & 0 \end{bmatrix}$$

13 Starting from $C = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $R = \begin{bmatrix} 2 & 4 \end{bmatrix}$ compute CR and RC and CRC and RCR.

14 Complete these 2 by 2 matrices to meet the requirements printed underneath:

$$\underset{\text{rank one}}{\begin{bmatrix} 3 & 6 \\ 5 & \end{bmatrix}} \qquad \underset{\text{orthogonal columns}}{\begin{bmatrix} 6 & \\ 7 & \end{bmatrix}} \qquad \underset{\text{rank 2}}{\begin{bmatrix} 2 & \\ 3 & 6 \end{bmatrix}} \qquad \underset{A^2 = I}{\begin{bmatrix} 3 & 4 \\ & -3 \end{bmatrix}}$$

15 Suppose $A = CR$ with independent columns in C and independent rows in R. Explain how each of these logical steps follows from $A = CR = (m \text{ by } r)\,(r \text{ by } n)$.

1. Every column of A is a combination of columns of C.
2. Every row of A is a combination of rows of R. What combination is row 1 ?
3. The number of columns of $C =$ the number of rows of R (needed for CR ?).
4. *Column rank equals row rank.* The number of independent columns of A equals the number of independent rows in A.

16 (a) The vectors $AB\boldsymbol{x}$ produce the column space of AB. Show why this vector $AB\boldsymbol{x}$ is also in the column space of A. (Is $AB\boldsymbol{x} = A\boldsymbol{y}$ for some vector $\boldsymbol{y}$?) Conclusion: The column space of A *contains* the column space of AB.

(b) Choose nonzero matrices A and B so the column space of AB contains only the zero vector. This is the smallest possible column space.

17 True or false, with a reason (not easy):

(a) If 3 by 3 matrices A and B have rank 1, then AB will always have rank 1.

(b) If 3 by 3 matrices A and B have rank 3, then AB will always have rank 3.

(c) Suppose $AB = BA$ for every 2 by 2 matrix B. Then $A = \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} = cI$ for some number c. Only those matrices A commute with every B.

18 Example 6 in this section mentioned a special case of the law $\boldsymbol{(AB)C = A(BC)}$.

$$A = C = \text{ exchange matrix } \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}.$$

(a) First compute AB (row exchange) and also BC (column exchange).

(b) Now compute the double exchanges: $(AB)C$ with rows first and $A(BC)$ with columns first. Verify that those double exchanges produce the same ABC.

19 Test the column-row multiplication in equation (5) to find AB and BA:

$$AB = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad BA = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

20 How many small multiplications for $(AB)C$ and $A(BC)$ if those matrices have sizes $ABC = (4 \times 3)\,(3 \times 2)\,(2 \times 1)$? That choice affects the operation count.

Thoughts on Chapter 1

Most textbooks don't have a place for the author's thoughts. But a lot of decisions go into starting a new textbook. This chapter has intentionally jumped right into the subject, with discussion of independence and rank. There are so many good ideas ahead, and they take time to absorb, so why not get started ? Here are two questions that influenced the writing.

What makes this subject easy ?	All the equations are linear.
What makes this subject hard ?	So many equations and unknowns and ideas.

Book examples are small size. But if we want the temperature at many points of an engine, there is an equation at every point: easily $n = 1000$ unknowns.

I believe the key is to work right away with matrices. $A\boldsymbol{x} = \boldsymbol{b}$ is a perfect format to accept problems of all sizes. The linearity is built into the symbols $A\boldsymbol{x}$ and the rule is $A(\boldsymbol{x} + \boldsymbol{y}) = A\boldsymbol{x} + A\boldsymbol{y}$. Each of the m equations in $A\boldsymbol{x} = \boldsymbol{b}$ represents a flat surface:

$2x + 5y - 4z = 6$ is a plane in three-dimensional space

$2x + 5y - 4z + 7w = 9$ is a 3D plane (hyperplane ?) in four-dimensional space

Linearity is on our side, but there is a serious problem in visualizing 10 planes meeting in 11-dimensional space. Hopefully they meet along a line: dimension $11 - 10 = 1$. An 11th plane should cut through that line at one point (which solves all 11 equations). What the textbook and the notation must do is to keep the counting simple

Here is what we expect for a random m by n matrix A:

$\boldsymbol{m < n}$ Many solutions or no solutions to the m equations $A\boldsymbol{x} = \boldsymbol{b}$

$\boldsymbol{m = n}$ Probably one solution to the n equations $A\boldsymbol{x} = \boldsymbol{b}$

$\boldsymbol{m > n}$ Probably no solution: too many equations with only n unknowns in $\boldsymbol{x}$

But this count is not necessarily what we get ! Columns of A can be combinations of previous columns: nothing new. An equation can be a combination of previous equations. **The rank $\boldsymbol{r}$ tells us the real size of our problem**, from independent columns and rows. The beautiful formula is $\boldsymbol{A = CR} = (m \times r)\,(r \times n)$: three matrices of rank r.

Notice: The columns of A that go into C must produce the matrix I inside R.

We end with the great associative law $\boldsymbol{(AB)\,C = A\,(BC)}$. Suppose C has 1 column:

AB has columns $A\boldsymbol{b}_1, \dots, A\boldsymbol{b}_n$ and then $(AB)\boldsymbol{c}$ equals $c_1 A\boldsymbol{b}_1 + \cdots + c_n A\boldsymbol{b}_n$.

$B\boldsymbol{c}$ has one column $c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n$ and $A(B\boldsymbol{c}) = A\,(c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n)$.

Linearity gives equality of those two sums. This proves $(AB)\,\boldsymbol{c} = A\,(B\boldsymbol{c})$.

The same is true for every column of C. Therefore $\boldsymbol{(AB)\,C = A\,(BC)}$.