1.4 Matrix Multiplication AB and 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 $\boldsymbol{i}$ of $\boldsymbol{A}$**) $\cdot$ (**column $\boldsymbol{j}$ of $\boldsymbol{B}$**).

3 By columns: **$\boldsymbol{A}$ times column $\boldsymbol{j}$ of $\boldsymbol{B}$ produces column $\boldsymbol{j}$ of $\boldsymbol{AB}$.**

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

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

We know how to multiply a matrix A times a column vector $\boldsymbol{x}$ or $\boldsymbol{b}$. This section moves to matrix-matrix multiplication: **a matrix $\boldsymbol{A}$ times a matrix $\boldsymbol{B}$**. The new rule builds on the old one, when the matrix B has columns $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_p$. We just multiply A times each of those p columns of B to find the p columns of AB.

Column $\boldsymbol{j}$ of $\boldsymbol{AB}$ equals $\boldsymbol{A}$ times column $\boldsymbol{j}$ of $\boldsymbol{B}$

$$\textbf{If } \boldsymbol{B} = \begin{bmatrix} \boldsymbol{b}_1 \cdots \boldsymbol{b}_p \end{bmatrix} \textbf{ then } \boldsymbol{AB} = \begin{bmatrix} \boldsymbol{Ab_1} \cdots \boldsymbol{Ab_p} \end{bmatrix} \tag{1}$$

To see that clearly, start with a 2 by 2 "exchange matrix" for B. So B has two columns $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$. We multiply A times each column to produce a column of AB:

$$A\boldsymbol{b}_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix} = \begin{bmatrix} \mathbf{2} \\ \mathbf{4} \end{bmatrix} \quad A\boldsymbol{b}_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{3} \end{bmatrix} \quad \boldsymbol{AB} = \begin{bmatrix} \mathbf{1} & \mathbf{2} \\ \mathbf{3} & \mathbf{4} \end{bmatrix}\begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{1} \\ \mathbf{4} & \mathbf{3} \end{bmatrix}$$

For this matrix B, the result of multiplying AB is to *exchange the columns of* A.

There is more to see when we multiply the same A by a full 2 by 2 matrix B:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} \mathbf{5} & \mathbf{6} \\ \mathbf{7} & \mathbf{8} \end{bmatrix} \text{ has } A\boldsymbol{b}_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} \mathbf{5} \\ \mathbf{7} \end{bmatrix} \text{ and } A\boldsymbol{b}_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} \mathbf{6} \\ \mathbf{8} \end{bmatrix}$$

Here is the point. We can multiply $A\boldsymbol{b}_1$ (matrix times vector) the *row way* or the *column way*. The row way uses dot products of $\boldsymbol{b}_1$ with *every row of* A:

Row way
Dot products

$$A\boldsymbol{b}_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 7 \end{bmatrix} = \begin{bmatrix} \text{row } 1 \cdot \boldsymbol{b}_1 \\ \text{row } 2 \cdot \boldsymbol{b}_1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 5 + 2 \cdot 7 \\ 3 \cdot 5 + 4 \cdot 7 \end{bmatrix} = \begin{bmatrix} \mathbf{19} \\ \mathbf{43} \end{bmatrix} \tag{2}$$

The column way uses a combination of the *columns* of A to find $A\boldsymbol{b}_1$. Same result:

Column way
Combine columns

$$A\boldsymbol{b}_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 7 \end{bmatrix} = \mathbf{5}\begin{bmatrix} 1 \\ 3 \end{bmatrix} + \mathbf{7}\begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 15 \end{bmatrix} + \begin{bmatrix} 14 \\ 28 \end{bmatrix} = \begin{bmatrix} \mathbf{19} \\ \mathbf{43} \end{bmatrix} \tag{3}$$

Both ways use the same 4 multiplications. With numbers like these, I think most people choose the row way. **To multiply $\boldsymbol{AB}$, take the dot product of each row of $\boldsymbol{A}$ with each column of $\boldsymbol{B}$**. When A has 2 rows and B has 2 columns, that means 4 dot products.

When A is m by n and B is n by p, then $\boldsymbol{AB}$ **is** $\boldsymbol{m}$ **by** $\boldsymbol{p}$. So we need mp dot products.

Row way
Rows of A
$$\boldsymbol{AB} = \begin{bmatrix} \text{row } 1 \cdot \text{col } 1 & \text{row } 1 \cdot \text{col } 2 \\ \text{row } 2 \cdot \text{col } 1 & \text{row } 2 \cdot \text{col } 2 \end{bmatrix} = \begin{bmatrix} \mathbf{19} & \mathbf{22} \\ \mathbf{43} & \mathbf{50} \end{bmatrix} \tag{4}$$

Now compute AB the **column way**: combinations of columns of A. This is a vector operation and it produces whole columns of AB. Equation (3) found the first column. Now we find 22 and 50 in the second column of AB from A times $\boldsymbol{b}_2$:

Column way
for $A\boldsymbol{b}_2$
$$A\boldsymbol{b}_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = \mathbf{6} \begin{bmatrix} 1 \\ 3 \end{bmatrix} + \mathbf{8} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 18 \end{bmatrix} + \begin{bmatrix} 16 \\ 32 \end{bmatrix} = \begin{bmatrix} \mathbf{22} \\ \mathbf{50} \end{bmatrix} \tag{5}$$

Equations (3) and (5) gave the same two columns of AB as equation (4). Both ways use *the same* 8 *multiplications*; only the order is different. To multiply an $\boldsymbol{m}$ **by** $\boldsymbol{n}$ matrix A times an $\boldsymbol{n}$ **by** $\boldsymbol{p}$ matrix B, we can count the small multiplications: AB is $\boldsymbol{m}$ **by** $\boldsymbol{p}$.

Row way $\boldsymbol{mp}$ dot products in AB, $\boldsymbol{n}$ multiplications each: $\boldsymbol{mnp}$ small multiplications

Column way $\boldsymbol{p}$ columns in AB, $\boldsymbol{mn}$ multiplications each: $\boldsymbol{mnp}$ small multiplications

The actual speed will depend on how the matrices are stored. I think column storage is usual. Please note that it is faster to move large pieces of a matrix from storage rather than individual numbers. In a big multiplication, matrix-matrix operations using BLAS 3 (Level 3 Basic Linear Algebra Subprograms) are the best. The comparison with Level 1 (*vector-vector*) and Level 2 (*matrix-vector*) is online at **netlib.org/blas/**.

So far we have used (row) $\cdot$ (column) dot products and (matrix) (column) $A\boldsymbol{b}_j$ in multiplying AB. The other two ways are (row) (matrix) and (column) (row), coming soon. All four ways use the same mnp multiplications in varying orders to find AB.

If A and B are 2 by 2, that means $n^3 = 8$ small multiplications for AB.[†] See below.

$\boldsymbol{AB}$ is usually different from $\boldsymbol{BA}$

For $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, AB exchanged the columns of A. **But $\boldsymbol{BA}$ exchanges the rows of $\boldsymbol{A}$!**

$$\boldsymbol{AB} = \begin{bmatrix} \mathbf{2} & \mathbf{1} \\ \mathbf{4} & \mathbf{3} \end{bmatrix} \qquad \boldsymbol{BA} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} \mathbf{3} & \mathbf{4} \\ \mathbf{1} & \mathbf{2} \end{bmatrix} \tag{6}$$

Matrix multiplication is not commutative. In general $\boldsymbol{BA} \neq \boldsymbol{AB}$. Multiply A on the left for row operations on A, and multiply on the right by B for column operations on A.

Question Why does squaring the exchange matrix give $B^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = I$?

[†] Strassen noticed that **7 multiplications** are enough for 2 by 2 matrices, at the cost of extra additions. For n by n matrices this reduces the multiplication count to $\boldsymbol{n^c}$, where $c = \log_2 7$ instead of the usual $c = \log_2 8 = 3$. Hard work has now reduced c even more. Certainly c cannot go below 2, because all of the n^2 entries in A and B must be used. Finding the smallest exponent c is an extremely tough unsolved problem.

AB times $C = A$ times BC

For matrix multiplication, **this associative law is true**. We are not willing to give up this extremely useful law. We can multiply AB first or we can multiply BC first.
The matrices stay in the order A, B, C and their sizes must be right for multiplication:

$$\boldsymbol{A} \textbf{ is } \boldsymbol{m\times n} \quad \boldsymbol{B} \textbf{ is } \boldsymbol{n\times p} \quad \boldsymbol{C} \textbf{ is } \boldsymbol{p\times q}. \text{ Then } \boldsymbol{AB} \textbf{ is } \boldsymbol{m\times p} \text{ and } \boldsymbol{(AB)C} \textbf{ is } \boldsymbol{m\times q}.$$

We can test the law using the exchange matrix B on the rows and the columns of A:

$$(\boldsymbol{BA})\boldsymbol{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{3} & \mathbf{4} \\ \mathbf{1} & \mathbf{2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{4} & \mathbf{3} \\ \mathbf{2} & \mathbf{1} \end{bmatrix}$$

$$\boldsymbol{B}(\boldsymbol{AB}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{2} & \mathbf{1} \\ \mathbf{4} & \mathbf{3} \end{bmatrix} = \begin{bmatrix} \mathbf{4} & \mathbf{3} \\ \mathbf{2} & \mathbf{1} \end{bmatrix}$$

So row operations on A can come *before or after* column operations on A.

Notice the meaning of $(AB)C = A(BC)$ when C is just a column vector $\boldsymbol{x}$. If that vector $\boldsymbol{x}$ has a single 1 in component j, then the associative law is $(AB)\boldsymbol{x} = A(B\boldsymbol{x})$. This tells us how to multiply matrices! The left side is **column $\boldsymbol{j}$ of $\boldsymbol{AB}$**. The right side is **$\boldsymbol{A}$ times column $\boldsymbol{j}$ of $\boldsymbol{B}$**. So their equality is exactly the rule for matrix multiplication that we saw in equation (1). It is simply the right rule.

Let me bring together the important facts about ABC and also A times $B + C$:

$$\boxed{\textbf{Associative} \quad (\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC}) \quad \text{and} \quad \textbf{Distributive} \quad \boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB}+\boldsymbol{AC}} \qquad (7)$$

Review of AB

Dot products (Row $\boldsymbol{i}$ of $\boldsymbol{A}$)$\cdot$(Col $\boldsymbol{j}$ of $\boldsymbol{B}$) $= (\boldsymbol{AB})_{\boldsymbol{ij}}$ = number in row i, col j of AB

Combine columns (Matrix $\boldsymbol{A}$) (Column $\boldsymbol{b}_j$ of $\boldsymbol{B}$) = vector in column j of AB

With numbers (the usual way), mp dot products produce the m by p matrix AB.
With vectors (the big picture), p combinations $A\boldsymbol{b}_j$ produce the p columns of AB.

For computing by hand, I would use the row way to find each number in AB. I visualize multiplication by columns: **The columns $\boldsymbol{Ab_j}$ in $\boldsymbol{AB}$ are combinations of columns of $\boldsymbol{A}$.**

Rank One Matrices and $A = CR$

All columns of a rank one matrix lie on the same line. That line is the column space of A. Examples in Section 1.3 pointed to a remarkable fact: *The rows also lie on a line*. When all the columns of A are in the same column direction, then all the rows of A are in the same row direction. Here is a new example of this extreme case: **rank $\boldsymbol{r} = \mathbf{1}$.**

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

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*? Another example: **Matrix of all 1's = (Column of 1's)** times **(Row of 1's)**.

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

Rank = 1

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 10 & 100 \\ 3 & 6 & 30 & 300 \\ 2 & 4 & 20 & 200 \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{3} \\ \mathbf{2} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{2} & \mathbf{10} & \mathbf{100} \end{bmatrix} = \boldsymbol{CR} \qquad (8)$$

The dot products (row of C) $\cdot$ (column of R) are small multiplications like 1 times 1. The last dot product is 2 times 100. We are following the dot product rule! This is multiplication of thin matrices CR. 12 small multiplications produce the 12 numbers in A.

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. Our next goal is to allow $\boldsymbol{r}$ **columns in $\boldsymbol{C}$** and to find $\boldsymbol{r}$ **rows in $\boldsymbol{R}$**. And to see $\boldsymbol{A} = \boldsymbol{CR}$. **That number $\boldsymbol{r}$ is the "rank" of $\boldsymbol{A}$.**

C Contains the First r Independent Columns of A

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 $\boldsymbol{A}$ and $\boldsymbol{C}$**. The n columns of A might be dependent. The r columns of C will surely be **independent**.

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

Those r independent columns in C combine to give all n columns in A.

$C\boldsymbol{x} = \mathbf{0}$ means that x_1(column 1 of C) + x_2(column 2 of C) + $\cdots$ = *zero vector*. With independent columns, $C\boldsymbol{x} = \mathbf{0}$ 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. C always has independent columns.

Example 2 $\quad A = \begin{bmatrix} \mathbf{2} & 6 & \mathbf{4} \\ \mathbf{4} & 12 & \mathbf{8} \\ \mathbf{1} & 3 & \mathbf{5} \end{bmatrix}$ leads to $C = \begin{bmatrix} \mathbf{2} & \mathbf{4} \\ \mathbf{4} & \mathbf{8} \\ \mathbf{1} & \mathbf{5} \end{bmatrix}$ **Rank $\boldsymbol{r} = \mathbf{2}$**

Columns 1 and 3 go into C. Column 2 is 3 times column 1: *not independent, not in C*.

Matrix Multiplication C times R

***R* tells how to produce all columns of *A* from the columns of *C*.** Then $A = CR$. 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 comes second in C, so the third column of R just has 0 and 1.

Notice I inside R Rank $r = 2$

$$A = CR \text{ is } \begin{bmatrix} 2 & 6 & 4 \\ 4 & 12 & 8 \\ 1 & 3 & 5 \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{4} \\ \mathbf{4} & \mathbf{8} \\ \mathbf{1} & \mathbf{5} \end{bmatrix} \begin{bmatrix} \mathbf{1} & ? & \mathbf{0} \\ \mathbf{0} & ? & \mathbf{1} \end{bmatrix}. \qquad (9)$$

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 is a dependent column. Column 2 of A is 3 times column 1, so *that number* 3 *goes into* R.

A is $m \times n$
C is $m \times r$
R is $r \times n$

$$\boxed{\boldsymbol{A = CR} \text{ is } \begin{bmatrix} 2 & 6 & 4 \\ 4 & 12 & 8 \\ 1 & 3 & 5 \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{4} \\ \mathbf{4} & \mathbf{8} \\ \mathbf{1} & \mathbf{5} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{3} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}} \qquad (10)$$

That example is typical of $A = CR$. We review the descriptions of C and R.

1. $\boldsymbol{C}$ contains a full set of r **independent columns** (chosen left to right) in A
2. $\boldsymbol{R} = [\, \boldsymbol{I} \quad \boldsymbol{F} \,]$ contains the **identity matrix $\boldsymbol{I}$** in the same r columns that held C.
3. **The dependent columns of $\boldsymbol{A}$ are combinations $\boldsymbol{CF}$** of the independent columns in C.

That matrix F goes into the other $n - r$ columns of $\boldsymbol{R} = [\, \boldsymbol{I} \quad \boldsymbol{F} \,]$. $A = CR$ becomes

$$\boldsymbol{A} = \boldsymbol{C}[\, \boldsymbol{I} \quad \boldsymbol{F} \,] = [\, \boldsymbol{C} \quad \boldsymbol{CF} \,] = [\, \textbf{indep cols of } \boldsymbol{A} \quad \textbf{dep cols of } \boldsymbol{A} \,] \ \textbf{(in correct order)}$$

$\boldsymbol{C}$ has the same column space as $\boldsymbol{A}$. $\boldsymbol{R}$ has the same row space as $\boldsymbol{A}$. Here $\boldsymbol{F} = \begin{bmatrix} \mathbf{-1} \\ \mathbf{2} \end{bmatrix}$

Example 3 of $\boldsymbol{A = CR}$ Rank 2

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{-1} \\ \mathbf{0} & \mathbf{1} & \mathbf{2} \end{bmatrix} \qquad (11)$$

When a column of $\boldsymbol{A}$ goes into $\boldsymbol{C}$, a column of $\boldsymbol{I}$ goes into $\boldsymbol{R}$.

$$\boxed{\textbf{Column } \boldsymbol{j} \textbf{ of } \boldsymbol{A} = \boldsymbol{C} \text{ times } \textbf{column } \boldsymbol{j} \textbf{ of } \boldsymbol{R}. \qquad \textbf{Row } \boldsymbol{i} \textbf{ of } \boldsymbol{A} = \textbf{row } \boldsymbol{i} \textbf{ of } \boldsymbol{C} \text{ times } \boldsymbol{R}.}$$

If all columns of A are independent, then $C = A$. What matrix is R? **Answer $\boldsymbol{R} = \boldsymbol{I}$.**

Chapter 1 finds $\boldsymbol{C}$ (independent columns of A) before R. Chapter 3 will find R first.

Here column 3 of A is the 2nd independent column in C. Then column 3 of R is $\begin{matrix}\mathbf{0}\\\mathbf{1}\end{matrix}$

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

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

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

1. The r columns of C are independent (they are chosen that way from A)
2. Every column of A is a combination of those r columns of C (this is $A = CR$)
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 the r rows of R (this is $A = CR$ by rows !)

How to Find the Matrix $\boldsymbol{R}$

Up to now you have had very little help in discovering the matrix R in $A = CR$. If you could tell that column 3 of this matrix A is a combination of columns 1 and 2, then the numbers x and y in that combination will go into column 3 of R:

$$\textbf{Example 4}\qquad \boldsymbol{A}=\begin{bmatrix}1&3&4\\2&4&2\\3&7&6\end{bmatrix}\qquad x\begin{bmatrix}1\\2\\3\end{bmatrix}+y\begin{bmatrix}3\\4\\7\end{bmatrix}\overset{?}{=}\begin{bmatrix}4\\2\\6\end{bmatrix}. \tag{12}$$

But even for this small matrix, we can't immediately see x and y. So we don't know the rank of A (2 or 3 **?**). There has to be a good way to discover x and y.

That good way is elimination. It will be the key algorithm in Chapter 2 for square matrices and again in Chapter 3 for all matrices. We want to introduce it now for this matrix.

The idea is to simplify A by "*row operations*". That will simplify the equations for x and y. We will **eliminate the 2 and 3** in column 1 of A. To do that, **subtract 2 times row 1 from row 2 of $\boldsymbol{A}$** and also **subtract 3 times row 1 from row 3**. The matrix A changes to B.

$$\boldsymbol{B}=\begin{bmatrix}1&3&4\\0&-2&-6\\0&-2&-6\end{bmatrix}\qquad x\begin{bmatrix}1\\0\\0\end{bmatrix}+y\begin{bmatrix}3\\-2\\-2\end{bmatrix}\overset{?}{=}\begin{bmatrix}4\\-6\\-6\end{bmatrix} \tag{13}$$

We only did what is legal. *Subtracting an equation from an equation leaves a new equation*. The new equation is $-2y = -6$, so we know $\boldsymbol{y = 3}$. Then if $\boldsymbol{x = -5}$ the top equation becomes $-5 + 9 = 4$, which is correct. The original equations (12) are solved by $-5, 3$:

$$\begin{matrix}x=-5\\y=+3\end{matrix}\qquad -\mathbf{5}\begin{bmatrix}1\\2\\3\end{bmatrix}+\mathbf{3}\begin{bmatrix}3\\4\\7\end{bmatrix}=\begin{bmatrix}4\\2\\6\end{bmatrix}\qquad \textbf{Column 3 of } \boldsymbol{A} \textbf{ is dependent}$$

So -5 and 3 are the numbers we needed in column 3 of R. All the ranks are $r = 2$:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 4 & 2 \\ 3 & 7 & 6 \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \mathbf{3} \\ \mathbf{2} & \mathbf{4} \\ \mathbf{3} & \mathbf{7} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{-5} \\ \mathbf{0} & \mathbf{1} & \mathbf{3} \end{bmatrix} = \boldsymbol{CR}. \tag{14}$$

There is more to see in this example. The elimination process that reduced A to B is called *row reduction*. I will complete it from B to U, to make the matrix even simpler. Just subtract row 2 of B from row 3 of B to see a **row of zeros in $\boldsymbol{U}$:**

$$\boldsymbol{A} \to \boldsymbol{B} = \begin{bmatrix} 1 & 3 & 4 \\ 0 & -2 & -6 \\ 0 & -2 & -6 \end{bmatrix} \to \boldsymbol{U} = \begin{bmatrix} \mathbf{1} & 3 & 4 \\ 0 & \mathbf{-2} & -6 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{matrix} \textbf{upper triangular} \\ \textbf{matrix } \boldsymbol{U} \end{matrix}. \tag{15}$$

That zero row is a clear signal: the row rank is also 2. Chapter 2 will stop with U. Chapter 3 will **eliminate upward** to produce more zeros. **We end up with $\boldsymbol{R_0}$ and $\boldsymbol{R}$**:

$$\boldsymbol{U} = \begin{bmatrix} 1 & 3 & 4 \\ 0 & -2 & -6 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} \mathbf{1} & 0 & \mathbf{-5} \\ 0 & \mathbf{1} & \mathbf{3} \\ 0 & 0 & 0 \end{bmatrix} = \boldsymbol{R_0}$$

All rows of R_0 are combinations of the original rows of A

That zero row of R_0 shows that A has rank $r = 2$

The 2 by 2 identity matrix shows that columns $1, 2$ of A are independent (in C)

Removing the zero row of $\boldsymbol{R_0}$ leaves the desired matrix $\boldsymbol{R}$ in $\boldsymbol{A = CR}$

Elimination in Chapter 3 will be a systematic way to find $\boldsymbol{R}$

Key facts	The r columns of C are a **basis** for the column space of A: **dimension $\boldsymbol{r}$**
$\boldsymbol{A = CR}$	The r rows of R are a **basis** for the row space of A: **dimension $\boldsymbol{r}$**

Those words **"basis"** and **"dimension"** will be properly defined later in Section 3.4.

Chapter 1 *starts with independent columns of* A, placed in C.
Chapter 3 *starts with the rows of* A*, and combines them into* R.

We are emphasizing CR because both matrices are so important. $\boldsymbol{C}$ contains r independent columns of A. $\boldsymbol{R}$ tells how to combine those columns to give all columns of A. (R contains I, because r columns of A are already in C.) Chapter 3 will produce R *directly from* A *by elimination*, the most used algorithm in computational mathematics.

$A = CR$ will be the key to a fundamental problem: *Solving linear equations* $A\boldsymbol{x} = \boldsymbol{b}$.

Columns of A times Rows of B . . . Columns of C times Rows of R

Before this chapter ends, I want to add this message. There is another way to multiply matrices (producing the same matrix AB or CR as always). This way is not so well known, but it is powerful. **The new way multiplies columns of A times rows of B.**

$$\boldsymbol{AB} = \begin{bmatrix} | & & | \\ \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \\ | & & | \end{bmatrix} \begin{bmatrix} — & \boldsymbol{b}_1^* & — \\ & \vdots & \\ — & \boldsymbol{b}_n^* & — \end{bmatrix} = \boldsymbol{a}_1\boldsymbol{b}_1^* + \boldsymbol{a}_2\boldsymbol{b}_2^* + \cdots + \boldsymbol{a}_n\boldsymbol{b}_n^*. \tag{16}$$

columns $\boldsymbol{a}_k$ rows $\boldsymbol{b}_k^*$ Add columns $\boldsymbol{a}_k$ times rows $\boldsymbol{b}_k^*$

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**). If A is an m by n matrix and B is an n by p matrix, then columns of A times rows of B adds up to the *same answer* AB as dot products of rows of A and columns of B.

AB involves the same mnp small multiplications but in a new order !

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

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

$$\textbf{Rank 1} + \textbf{Rank 1} = \begin{bmatrix} 7 & 8 & 9 \\ 14 & 16 & 18 \\ 21 & 24 & 27 \end{bmatrix} + \begin{bmatrix} 40 & 44 & 48 \\ 50 & 55 & 60 \\ 60 & 66 & 72 \end{bmatrix} = \begin{bmatrix} \mathbf{47} & \mathbf{52} & \mathbf{57} \\ \mathbf{64} & \mathbf{71} & \mathbf{78} \\ \mathbf{81} & \mathbf{90} & \mathbf{99} \end{bmatrix} = \boldsymbol{AB}$$

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 uses different words. *AB has no inverse matrix : it is not invertible.* And in Chapter 5 : *The determinant of AB is zero.*

Note about the matrix R

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.**

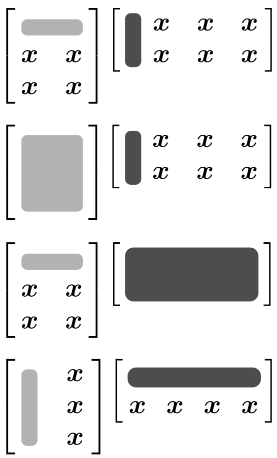
Example 5 $$A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \mathbf{3}\boldsymbol{a}_1 + \mathbf{4}\boldsymbol{a}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 \end{bmatrix} \begin{bmatrix} \mathbf{1} & 0 & \mathbf{3} \\ 0 & \mathbf{1} & \mathbf{4} \end{bmatrix} = \boldsymbol{CR}.$$

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 5 :

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

(3 by 2) (2 by 4) = (3 by 4) **Four Ways to Multiply $\boldsymbol{AB = C}$**



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

$\boldsymbol{A}$ times (**Column $\boldsymbol{k}$** of B) = **Column $\boldsymbol{k}$ of $\boldsymbol{C}$**
$k = 1$ to 4 **4 columns**

(**Row $\boldsymbol{i}$** of A) times $\boldsymbol{B}$ = **Row $\boldsymbol{i}$ of $\boldsymbol{C}$**
$i = 1$ to 3 **3 rows**

(**Column $\boldsymbol{j}$** of A) (**Row $\boldsymbol{j}$** of B) = **Rank 1 Matrix**
$j = 1$ to 2 **2 matrices**

Dot product way, Column way, Row way, Columns times rows

Problem Set 1.4

1 Rewrite this four-way table for $AB = C$ 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 $\boldsymbol{a} \neq \boldsymbol{0}$, what are C and R ?

3 Multiply A times B (3 examples) using *dot products* : (each row) $\cdot$ (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) $\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ $\qquad$ (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 $B = CR$?

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

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, when you go 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 :

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

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 $A = cI$ commute with every B.

18 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 matrix multiplication in equation (16) 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 $ABC = (4 \times 3)\,(3 \times 2)\,(2 \times 1)$? The two counts are different.

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:

$\boldsymbol{m < n}$ Probably many 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 multiply 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 $A\boldsymbol{b}_1, \ldots, A\boldsymbol{b}_n$ 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 then $A(B\boldsymbol{c}) = A\,(c_1\boldsymbol{b}_1 + \cdots + c_n\boldsymbol{b}_n)$.

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 $\boldsymbol{(AB)\,C = A\,(BC)}$.