1 Matrix-Vector Multiplication: \( \mathbf{A} \mathbf{x} \) by \( \rightarrow \downarrow \)

The first time one encounters matrix multiplication, it may seem arbitrary. First we will feed this impression by laying out a way of multiplying \( m \times n \) matrix \( \mathbf{A} \) by \( n \times 1 \) column vector \( \mathbf{x} \). But then we will look at some examples which should make this product seem more natural.

Let

\[
\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \quad \text{with} \quad \mathbf{x} = \begin{bmatrix} 20 \\ 30 \\ 40 \\ 50 \end{bmatrix}
\]

so that

\[
\mathbf{A} \mathbf{x} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \begin{bmatrix} 20 \\ 30 \\ 40 \\ 50 \end{bmatrix} = \begin{bmatrix} 400 \\ 960 \\ 1520 \end{bmatrix},
\]

which says that the product \( \mathbf{A} \mathbf{x} \) is a vector with the same number of rows as \( \mathbf{A} \).

The first entry of \( \mathbf{A} \mathbf{x} \) is the number

\[
\]

One can think of this as obtained by scanning the first row of \( \mathbf{A} \) horizontally from left to right and simultaneously scanning down \( \mathbf{x} \), matching corresponding entries to form products to be added at the end.

The second entry of \( \mathbf{A} \mathbf{x} \) is formed the same way, except that the second row of \( \mathbf{A} \) is used:

\[
\]

And the third entry, using the third row of \( \mathbf{A} \):

\[
\]

Accordingly,

\[
\mathbf{A} \mathbf{x} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \begin{bmatrix} 20 \\ 30 \\ 40 \\ 50 \end{bmatrix} = \begin{bmatrix} 400 \\ 960 \\ 1520 \end{bmatrix}.
\]
Make sure you agree with the following results when you use this method:

(a) \[
\begin{bmatrix}
2 & -3 \\
5 & 4 \\
\end{bmatrix}
\begin{bmatrix}
6 \\
4 \\
\end{bmatrix} =
\begin{bmatrix}
0 \\
46 \\
\end{bmatrix}
\]

(b) \[
\begin{bmatrix}
2 & -3 \\
-4 & 6 \\
\end{bmatrix}
\begin{bmatrix}
6 \\
4 \\
\end{bmatrix} =
\begin{bmatrix}
0 \\
0 \\
\end{bmatrix}
\]

(c) \[
\begin{bmatrix}
4 & 2 & -3 \\
-4 & 3 & 5 \\
\end{bmatrix}
\begin{bmatrix}
8 \\
-5 \\
\end{bmatrix}
\]
is undefined.

(d) \[
\begin{bmatrix}
2 & -3 \\
-4 & 6 \\
\end{bmatrix}
\begin{bmatrix}
3 \\
2 \\
\end{bmatrix}
\]
is also undefined.

(e) \[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
2 \\
6 \\
8 \\
\end{bmatrix} =
\begin{bmatrix}
2 \\
6 \\
8 \\
\end{bmatrix}
\]

(f) \[
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
p \\
q \\
\end{bmatrix} =
\begin{bmatrix}
p \\
q \\
\end{bmatrix}
\]

(g) \[
\begin{bmatrix}
4 & 2 & -3 \\
-4 & 3 & 5 \\
\end{bmatrix}
\begin{bmatrix}
3 \\
-1 \\
2 \\
\end{bmatrix} =
\begin{bmatrix}
4 \\
-5 \\
\end{bmatrix}
\]

(h) \[
\begin{bmatrix}
4 & 2 & -3 \\
-4 & 3 & 5 \\
\end{bmatrix}
\begin{bmatrix}
a \\
b \\
c \\
\end{bmatrix} =
\begin{bmatrix}
4a + 2b - 3c \\
-4a + 3b + 5c \\
\end{bmatrix}
\]

Example (h) above suggests the following “matrification” of a system of linear equations:

\[
\begin{align*}
3x - 5y - 2z &= 17 \\
5x - 6y - 2z &= 24 \\
-2x + 4y + 2z &= -12 \\
\end{align*}
\]

or

\[
\begin{bmatrix}
3x - 5y - 2z \\
5x - 6y - 2z \\
-2x + 4y + 2z \\
\end{bmatrix} =
\begin{bmatrix}
17 \\
24 \\
-12 \\
\end{bmatrix}
\]

or

\[
\begin{bmatrix}
3 & -5 & -2 \\
5 & -6 & -2 \\
-2 & 4 & 2 \\
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
\end{bmatrix} =
\begin{bmatrix}
17 \\
24 \\
-12 \\
\end{bmatrix}
\]

or

\[A\vec{x} = \vec{b}, \text{ where } A = \begin{bmatrix}
3 & -5 & -2 \\
5 & -6 & -2 \\
-2 & 4 & 2 \\
\end{bmatrix}, \vec{x} = \begin{bmatrix}
x \\
y \\
z \\
\end{bmatrix}, \text{ and } \vec{b} = \begin{bmatrix}
17 \\
24 \\
-12 \\
\end{bmatrix}.
\]

Matrix \( A \) is known as the coefficient matrix of the system. \( \vec{b} \) is sometimes called the vector of right-hand sides.
Examples (e) and (f) show us the $3 \times 3$ and $2 \times 2$ instances of the identity matrix $I$. In a $3 \times 3$ context,

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

while in a $2 \times 2$ context,

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
2 A Rotation in \( \mathbb{R}^2 \) via Matrix Multiplication

The following figure shows vector \( \vec{b} \) obtained from vector \( \vec{a} \) by rotating \( \vec{a} \) through a positive angle \( \theta \):

![Diagram showing vector \( \vec{a} \) rotated to \( \vec{b} \) through angle \( \theta \)]

Now, the polar form of \( \vec{a} \) is given by

\[
\vec{a} = ||\vec{a}|| \begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \end{bmatrix} = \begin{bmatrix} ||\vec{a}|| \cos(\alpha) \\ ||\vec{a}|| \sin(\alpha) \end{bmatrix}
\]

And, since \( ||\vec{b}|| = ||\vec{a}|| \) and \( \vec{b} \) makes angle \( \theta + \alpha \) with the positive \( x \)-axis, the polar form of \( \vec{b} \) can be expressed as follows:

\[
\vec{b} = ||\vec{a}|| \begin{bmatrix} \cos(\theta + \alpha) \\ \sin(\theta + \alpha) \end{bmatrix} = ||\vec{a}|| \begin{bmatrix} \cos(\theta) \cos(\alpha) - \sin(\theta) \sin(\alpha) \\ \sin(\theta) \cos(\alpha) + \cos(\theta) \sin(\alpha) \end{bmatrix},
\]

so that

\[
\vec{b} = \begin{bmatrix} \cos(\theta) ||\vec{a}|| \cos(\alpha) - \sin(\theta) ||\vec{a}|| \sin(\alpha) \\ \sin(\theta) ||\vec{a}|| \cos(\alpha) + \cos(\theta) ||\vec{a}|| \sin(\alpha) \end{bmatrix} = \begin{bmatrix} \cos(\theta) - \sin(\theta) \\ \sin(\theta) \cos(\theta) \end{bmatrix} \begin{bmatrix} ||\vec{a}|| \cos(\alpha) \\ ||\vec{a}|| \sin(\alpha) \end{bmatrix}
\]

or

\[
\vec{b} = \begin{bmatrix} \cos(\theta) - \sin(\theta) \\ \sin(\theta) \cos(\theta) \end{bmatrix} \vec{a}.
\]

This formula is derived at this point to lend credibility to the idea of the product of matrix \( A \) times vector \( \vec{x} \).
3 \( A\vec{c} \) as a Linear Combination of Columns of \( A \)

Let \( A \) be a matrix and let \( \vec{c} \) be a vector such that \( A\vec{c} \) is defined.

It turns out that the vector \( A\vec{c} \) can always be looked upon as a linear combination of the columns of \( A \) with the entries of \( \vec{c} \) providing the coefficients. We can write down a proof of this later. For now, an example:

Let
\[
A = \begin{bmatrix}
2 & 4 & 1 & 0 \\
1 & 0 & 3 & -2 \\
2 & 1 & 7 & 5
\end{bmatrix}
\quad \text{with} \quad \vec{c} = \begin{bmatrix}
a \\
b \\
c \\
d
\end{bmatrix}
\]

then
\[
A\vec{c} = \begin{bmatrix}
2 & 4 & 1 & 0 \\
1 & 0 & 3 & -2 \\
2 & 1 & 7 & 5
\end{bmatrix}
\begin{bmatrix}
a \\
b \\
c \\
d
\end{bmatrix}
= \begin{bmatrix}
2a + 4b + 1c + 0d \\
1a + 0b + 3c - 2d \\
2a + 1b + 7c + 5d
\end{bmatrix}
= \begin{bmatrix}
2a \\
1a \\
2a
\end{bmatrix} + \begin{bmatrix}
4b \\
0b \\
1b
\end{bmatrix} + \begin{bmatrix}
1c \\
3c \\
7c
\end{bmatrix} + \begin{bmatrix}
0d \\
-2d \\
5d
\end{bmatrix}
\]

or
\[
A\vec{c} = a \begin{bmatrix}
2 \\
1 \\
2
\end{bmatrix} + b \begin{bmatrix}
4 \\
0 \\
1
\end{bmatrix} + c \begin{bmatrix}
1 \\
3 \\
7
\end{bmatrix} + d \begin{bmatrix}
0 \\
-2 \\
5
\end{bmatrix}
\]

4 \( A\vec{x} \) Entry as a Sum of Indexed Expressions

Here is yet a third way to look at the matrix-vector product \( A\vec{x} \).

We begin by introducing the traditional indexing schemes: for \( \vec{x} \) in \( \mathbb{R}^n \), we have the entries, top down, \( x_1, x_2, x_3, \) and so on, down to \( x_n \) at the bottom. Hence, if, in \( \mathbb{R}^3 \),
\[
z = \begin{bmatrix}
1 \\
3 \\
7
\end{bmatrix}, \text{ then } z_2 = 3 \text{ and } z_3 = 7.
\]

For \( \vec{x} \) and \( \vec{y} \) in \( \mathbb{R}^n \), we have
\[
\vec{x} \cdot \vec{y} = \sum_{k=1}^{n} x_k y_k \quad \text{and} \quad ||\vec{x}||^2 = \vec{x} \cdot \vec{x} = \sum_{t=1}^{n} x_t x_t
\]
For $m \times n$ matrix $A$, the indexing agreement goes like this:

$$A_{rc}$$

is the entry of $A$ in row $r$ and column $c$. Hence, if

$$A = \begin{bmatrix} 2 & 4 & 1 & 0 \\ 1 & 0 & 3 & -2 \\ 2 & 1 & 7 & 5 \end{bmatrix},$$

then $4 = A_{12}$, $A_{21} = 1$, $A_{33} = 7$, and $-2 = A_{24}$.

The thing here is that, if $A$ is $m \times n$ and $\vec{x}$ is an $n \times 1$ column vector in $\mathbb{R}^n$, then $A\vec{x}$ is an $m \times 1$ column vector in $\mathbb{R}^m$ and entry $r$ of $A\vec{x}$ is given by

$$(A\vec{x})_r = \sum_{k=1}^{n} A_{rk}x_k,$$

scanning across row $r$ of $A$ while scanning down vector $\vec{x}$.

We are thus left with three views of the matrix-vector product $A\vec{x}$:

(a) The $\rightarrow \downarrow$ method, which is the method you will most likely use.

(b) The idea that $A\vec{x}$ is a linear combination of columns of $A$ with entries of $\vec{x}$ furnishing the coefficients for the linear combination.

(c) The index-formula idea:

$$(A\vec{x})_k = \sum_{j=1}^{n} A_{kj}x_j,$$

For identity matrices, the entry $I_{rc}$ is traditionally written down as Kronecker’s $\delta$:

$$\delta_{rc} = \begin{cases} 1 & \text{if } r = c \\ 0 & \text{if } r \neq c \end{cases}$$
5 Matrix-Matrix Multiplication in Terms of Matrix-Vector Products

Again we present an operation somewhat arbitrarily, then justify it. If \( A \) and \( B \) are matrices, we form the matrix product \( AB \) by having \( A \) operate on the columns of \( B \) individually:

Let \( B \) have \( p \) columns \( \bar{B}_1, \ldots, \bar{B}_q \). Then \( AB \) is the matrix whose columns, in order, are \( A\bar{B}_1, \ldots, A\bar{B}_q \).

So, for example, if

\[
A = \begin{bmatrix} 5 & 4 & 0 \\ 2 & 4 & 3 \\ -1 & 6 & 3 \end{bmatrix} \quad \text{with} \quad B = \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 3 & 2 \end{bmatrix}
\]

Then \( B \) has columns

\[
\bar{B}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \text{and} \quad \bar{B}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix},
\]

so that

\[
A\bar{B}_1 = \begin{bmatrix} 9 \\ 15 \\ 14 \end{bmatrix} \quad \text{and} \quad A\bar{B}_2 = \begin{bmatrix} 14 \\ 14 \\ 10 \end{bmatrix},
\]

so that finally

\[
AB = \begin{bmatrix} 9 & 14 \\ 15 & 14 \\ 14 & 10 \end{bmatrix}.
\]

Certainly, if the columns of \( B \) aren’t the right size, then \( AB \) won’t be defined: the number of columns of \( A \) must be the same as the number of rows of \( B \).

6 \( AB \) by indices and \( AB \) by \( \rightarrow \downarrow \)

It’s necessary to have an index formula for \( AB \), that is, a formula for \( (AB)_{rc} \), the \( AB \) entry in row \( r \) and column \( c \).

And to that end, we note that \( (AB)_{rc} \) should be entry number \( r \) of the matrix-vector product \( A\bar{B}_c \).
Thus
\[(AB)_{rc} = (A\bar{B}_c)_r = \sum_{k=1}^{n} A_{rk}(\bar{B}_c)_k = \sum_{k=1}^{n} A_{rk}B_{kc},\]

where \(A\) is \(m \times n\) and \(B\) is \(n \times p\). Note that this last sum scans across row \(r\) of \(A\) and down column \(c\) of \(B\). This is why most folks compute matrix products using the \(\rightarrow \downarrow\) approach this index formula suggests.

Check that you agree with the following examples:

(a) \[
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9
\end{bmatrix}
\begin{bmatrix}
2 & -3 & 1 \\
-4 & 6 & -2 \\
2 & -3 & 1
\end{bmatrix}
= \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\]
\(AB = 0\), though neither \(A\) nor \(B\) is zero.

(b) \[
\begin{bmatrix}
2 & -3 & 1 \\
-4 & 6 & -2 \\
2 & -3 & 1
\end{bmatrix}
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9
\end{bmatrix}
= \begin{bmatrix}
-3 & -3 & -3 \\
6 & 6 & 6 \\
-3 & -3 & -3
\end{bmatrix}
\]
\(AB \neq BA\) sometimes.

(c) \[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9
\end{bmatrix}
= \begin{bmatrix}
1 & 2 & 3 \\
4 & 5 & 6 \\
7 & 8 & 9
\end{bmatrix}
\]
Identity matrix.

(d) \[
\begin{bmatrix}
1 & 2 \\
3 & 4
\end{bmatrix}
\begin{bmatrix}
4 & -2 \\
-3 & 1
\end{bmatrix}
= \begin{bmatrix}
-2 & 0 \\
0 & -2
\end{bmatrix}
\]

(e) \[
\begin{bmatrix}
6 & 5 & 4
\end{bmatrix}
\begin{bmatrix}
2
\end{bmatrix}
= \begin{bmatrix}
28
\end{bmatrix}
\]
\(f) \[
\begin{bmatrix}
1 \\
2 \\
3
\end{bmatrix}
\begin{bmatrix}
6 & 5 & 4
\end{bmatrix}
= \begin{bmatrix}
6 & 5 & 4 \\
12 & 10 & 8 \\
18 & 15 & 12
\end{bmatrix}
\]

(g) \[
\begin{bmatrix}
\cos(\alpha) & -\sin(\alpha) \\
\sin(\alpha) & \cos(\alpha)
\end{bmatrix}
\begin{bmatrix}
\cos(\beta) & -\sin(\beta) \\
\sin(\beta) & \cos(\beta)
\end{bmatrix}
=
\begin{bmatrix}
\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta) - \sin(\alpha)\cos(\beta) \\
\sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) & -\sin(\alpha)\sin(\beta) + \cos(\alpha)\cos(\beta)
\end{bmatrix}
=
\begin{bmatrix}
\cos(\alpha + \beta) & -\sin(\alpha + \beta) \\
\sin(\alpha + \beta) & \cos(\alpha + \beta)
\end{bmatrix}
\]
7 Transpose of a Matrix and Vector

If \( A \) is a matrix, we form its transpose \( A^T \) by making all the rows of \( A \) into columns of \( A^T \):

If 
\[
A = \begin{bmatrix}
1 & 2 \\
3 & 4 \\
5 & 6
\end{bmatrix},
\text{ then } A^T = \begin{bmatrix}
1 & 3 & 5 \\
2 & 4 & 6
\end{bmatrix}.
\]

We have 
\[
\begin{bmatrix}
6 \\
5 \\
4
\end{bmatrix}^T = \begin{bmatrix}
6 & 5 & 4
\end{bmatrix}.
\]

Note that the dot product of two vectors can be written in terms of multiplication and transpose:
\[
\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y}.
\]

If \( M = \begin{bmatrix}
\cos(\theta) & -\sin(\theta) \\
\sin(\theta) & \cos(\theta)
\end{bmatrix} \), then \( MM^T = ? \)

8 Adding Matrices and Multiplying a Matrix by a Number

Two matrix operations, addition and multiplying by a number, go just as for vectors:

(a) The sum of two matrices \( A \) and \( B \) of the same size, shape, and orientation is a matrix of the same size, shape and orientation, each entry of which is the sum of the corresponding entries of \( A \) and \( B \):

\[
(A + B)_{rc} = A_{rc} + B_{rc}
\]

for each row \( r \) and column \( c \).

(b) We multiply matrix \( A \) by number \( k \) by multiplying all entries of \( A \) by \( k \):

\[
(kA)_{rc} = k(A_{rc})
\]

So, for example, we have
\[
3 \begin{bmatrix}
1 & 2 \\
3 & 4
\end{bmatrix} + 5 \begin{bmatrix}
3 & -2 \\
0 & 10
\end{bmatrix} = \begin{bmatrix}
18 & -4 \\
9 & 62
\end{bmatrix}
\]
9 Associativity of Matrix Multiplication and other Algebraic Laws

We can use the index notation to prove some laws of matrix algebra: for matrices $A$, $B$, and $C$ with numbers $\alpha$ and $\beta$ we have

(A) $\alpha(A + B) = \alpha A + \alpha B$
(B) $(\alpha + \beta)C = \alpha C + \beta C$
(C) $(A + B)C = AC + BC$
(D) $C(A + B) = CA + CB$
(E) $C(\alpha A + \beta B) = \alpha CA + \beta CB$
(F) $(\alpha A + \beta B)C = \alpha AC + \beta BC$
(G) $\alpha(AB) = (\alpha A)B = A(\alpha B)$
(H) $A(BC) = (AB)C$

This is a few more formulas than we can get worked up about proving, but we should look at at least a couple.

The following “indexing” proofs make use of some finite-sum identities. Letting $x_1, \ldots, x_Q$ and $y_1, \ldots, y_Q$ be real numbers, we have:

(i) $\alpha \sum_{k=1}^{Q} x_k = \sum_{k=1}^{Q} \alpha x_k$
(ii) $\sum_{k=1}^{Q} x_i + \sum_{k=1}^{Q} y_i = \sum_{k=1}^{Q} (x_i + y_i)$
(iii) If $M$ is a $P \times Q$ matrix, then

$$\sum_{i=1}^{P} \left( \sum_{j=1}^{Q} M_{ij} \right) = \sum_{j=1}^{Q} \left( \sum_{i=1}^{P} M_{ij} \right),$$

so that summing the row sums has the same result as summing the column sums.
Here is a proof of the “Super-Distributivity” law

\[ C(\alpha A + \beta B) = \alpha CA + \beta CB \]

\[ LHS_{rc} = (C(\alpha A + \beta B))_{rc} \]
\[ = \sum_{k=1}^{n} C_{rk}(\alpha A + \beta B)_{kc} \]
\[ = \sum_{k=1}^{n} C_{rk}(\alpha A_{kc} + \beta B_{kc}) \]
\[ = \sum_{k=1}^{n} (C_{rk}\alpha A_{kc} + C_{rk}\beta B_{kc}) \]
\[ = \sum_{k=1}^{n} C_{rk}\alpha A_{kc} + \sum_{k=1}^{n} C_{rk}\beta B_{kc} \]
\[ = \alpha \left( \sum_{k=1}^{n} C_{rk}A_{kc} \right) + \beta \left( \sum_{k=1}^{n} C_{rk}B_{kc} \right) \]
\[ = \alpha (CA)_{rc} + \beta (CB)_{rc} \]
\[ = (\alpha CA)_{rc} + (\beta CB)_{rc} \]
\[ = (\alpha CA + \beta CB)_{rc} \]
\[ = RHS_{rc} \]

This completes the proof of law (E). Note that we have to assume that \( C \) is \( m \times n \), while \( A \) and \( B \) are \( n \times p \).
For the Associative Law of Matrix Multiplication (H) above, let's assume that $A$ is $m \times n$, $B$ is $n \times p$, and $C$ is $p \times q$. This means that $BC$ is $n \times q$, so that $A(BC)$ is $m \times q$. Furthermore, $AB$ is $m \times p$, so that $(AB)C$ is $m \times q$.

As before, we prove that the entries in row $r$ and column $c$ are identical for $A(BC)$ and $(AB)C$.

\[
(A(BC))_{rc} = \sum_{k=1}^{n} A_{rk} (BC)_{kc} \quad \text{(multiply $A$ and $BC$)}
\]

\[
= \sum_{k=1}^{n} A_{rk} \left( \sum_{\ell=1}^{p} B_{k\ell} C_{\ell c} \right) \quad \text{(multiply $B$ and $C$)}
\]

\[
= \sum_{k=1}^{n} \left( \sum_{\ell=1}^{p} A_{rk} B_{k\ell} C_{\ell c} \right) \quad \text{(multiply $A_{rk}$ though the inner sum)}
\]

\[
= \sum_{\ell=1}^{p} \left( \sum_{k=1}^{n} A_{rk} B_{k\ell} C_{\ell c} \right) \quad \text{(change order of summation)}
\]

\[
= \sum_{\ell=1}^{p} (AB)_{r\ell} C_{\ell c} \quad \text{(factor out $C_{\ell c}$)}
\]

\[
= ((AB)C)_{rc}
\]

This completes a proof of the Associative Law (H). We need to use it in the very next section of the text.
10 Cramer’s Rule, $2 \times 2$—Matrix Version

A system of two linear equations in two unknowns $x$ and $y$ can be considered to be a pair of straight-line equations:

\[
\begin{align*}
ax + by & = E \\
cx + dy & = F
\end{align*}
\]

Think of $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, with $\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} E \\ F \end{bmatrix}$ so the original system can be thought of as

\[A\vec{x} = \vec{y}\]

Also do the trick of forming $B = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ and then multiply as follows:

\[
B(A\vec{x}) = B\vec{y} \\
(BA)\vec{x} = B\vec{y}
\]

Now

\[
BA = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = (ad - bc) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = (ad - bc)I,
\]

so that

\[
B(A\vec{x}) = B\vec{y} \\
(BA)\vec{x} = B\vec{y} \\
(ad - bc)I\vec{x} = B\vec{y} \\
(ad - bc)\vec{x} = B\vec{y}.
\]

You can see that, as long as $(ad - bc) \neq 0$, we can get at $\vec{x}$ by numerically dividing through by $(ad - bc)$:

\[
\vec{x} = \frac{1}{(ad - bc)}B\vec{y} = \frac{1}{(ad - bc)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} E \\ F \end{bmatrix}.
\]

This last formula is $2 \times 2$ Matrix Cramer’s Rule.

The tricky matrix $B$ is sometimes called the adjugate of $A$, denoted $\text{adj}(A)$.

The expression $ad - bc$ is known as the determinant of the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:

\[
\det(A) = ad - bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.
\]
The crossing point of the lines $2x + 5y = 28$ and $3x + 4y = -35$ can be found by solving the system

$$
\begin{bmatrix}
2 & 5 \\
3 & 4
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix}
= 
\begin{bmatrix}
28 \\
-35
\end{bmatrix}.
$$

Thinking of $A = \begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix}$, we check the the determinant:

$$
\det(A) = \begin{vmatrix} 2 & 5 \\ 3 & 4 \end{vmatrix} = (2)(4) - (5)(3) = -7 \neq 0,
$$

so that

$$
\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{-7} \begin{bmatrix} 4 & -5 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 28 \\ -35 \end{bmatrix} = \frac{1}{-7} \begin{bmatrix} (4)(28) + (-5)(-35) \\ (-3)(28) + (2)(-35) \end{bmatrix} = \frac{1}{-7} \begin{bmatrix} 287 \\ -154 \end{bmatrix} = \begin{bmatrix} -41 \\ 22 \end{bmatrix},
$$

so that the lines cross at $(-41, 22)$.

Here is a system for which Matrix Cramer will not work:

$$
3x - 10y = 51 \quad \text{with} \quad -6x + 20y = \pi.
$$

We see that the determinant of the coefficient matrix evaluates to zero, so Matrix Cramer is out of the picture. If we fall back on GJ, we see:

$$
\begin{bmatrix} 3 & -10 \\ -6 & 20 \end{bmatrix} \begin{bmatrix} 51 \\ \pi \end{bmatrix} \quad \text{goes to} \quad \begin{bmatrix} 1 & -\frac{10}{3} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 17 \\ 102 + \pi \end{bmatrix}
$$

so that the system has no solutions at all – the straight lines involved are parallel. We could have scoped this out by noticing that the two lines have parallel normal vectors.

Here is another system for which Matrix Cramer will not work:

$$
2x + 11y = 128 \quad \text{with} \quad 4x + 22y = 256
$$

The determinant of the coefficients is zero and the RREF of the system is

$$
\begin{bmatrix} 1 & \frac{11}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 64 \\ 0 \end{bmatrix},
$$

so that there are infinitely many solutions:

$$
\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 64 \\ 0 \end{bmatrix} + t \begin{bmatrix} -11 \\ 2 \end{bmatrix}.
$$
11 The Range of the Function $f(\vec{x}) = A\vec{x}$

Recall from elementary algebra that the *range* of a function $f$ is the set of all possible values of the function. That is, $y$ is in the range of $f$ if there is an $x$ in the domain of $f$ such that $f(x) = y$. So, for example, $1/2$ is in the range of the sine function, because $1/2 = \sin(5\pi/6)$, while $2$ is not in the range of the sine, because for all $x$, $|\sin(x)| \leq 1$.

If $f$ is of the form $f(\vec{x}) = A\vec{x}$, then determining whether $\vec{y}$ is in the range of $f$ boils down to deciding whether $A\vec{x} = \vec{y}$ is a system with a solution.

Let

$$A = \begin{bmatrix} -8 & -24 & 3 \\ -3 & -9 & 1 \\ 10 & 30 & -3 \end{bmatrix}.$$ 

You may have worked problem of solving the system

$$A\vec{x} = \begin{bmatrix} 63 \\ 23 \\ -75 \end{bmatrix}.$$ 

This system turns out to have many solutions, so the vector $\vec{y} = \begin{bmatrix} 63 \\ 23 \\ -75 \end{bmatrix}$ does turn out to be in the range of the function $f$ given by

$$f(\vec{x}) = A\vec{x}.$$ 

We will see in a moment that not all vectors in $\mathbb{R}^3$ belong to this range, however. We can say that we want to find all vectors $\vec{b}$ such that $A\vec{x} = \vec{b}$ has a solution, that is, such that

$$
\begin{align*}
-8x - 24y + 3z &= b_1 \\
-3x - 9y + z &= b_2 \\
10x + 30y - 3z &= b_3
\end{align*}
$$

has a solution. The augmented matrix for this system looks like this:

$$
\begin{bmatrix}
-8 & -24 & 3 & | & b_1 \\
-3 & -9 & 1 & | & b_2 \\
10 & 30 & -3 & | & b_3
\end{bmatrix}.
$$

A few steps into GJ yields a cumbersome mess of expressions on the right-hand side. We feel the need of a trick to make life a bit easier. Here's one: we have columns for the
coefficients of \( x, y, \) and \( z \). Let’s put in some columns for the coefficients of \( b_1, b_2, \) and \( b_3 \):

\[
\begin{bmatrix}
-8 & -24 & 3 & | & 1 & 0 & 0 \\
-3 & -9 & 1 & | & 0 & 1 & 0 \\
10 & 30 & -3 & | & 0 & 0 & 1
\end{bmatrix}.
\]

We do EROs to this matrix until we have the RREF of matrix \( A \) on the left side of the divider:

\[
\begin{bmatrix}
1 & 3 & -\frac{3}{8} & | & -\frac{1}{8} & 0 & 0 \\
0 & 0 & -\frac{1}{8} & | & -\frac{3}{8} & 1 & 0 \\
0 & 0 & \frac{3}{4} & | & \frac{5}{4} & 0 & 1
\end{bmatrix}
\begin{bmatrix}
1 & 3 & 0 & | & 1 & -3 & 0 \\
0 & 0 & 1 & | & 3 & -8 & 0 \\
0 & 0 & 0 & | & -1 & 6 & 1
\end{bmatrix}.
\]

Taking this super-augmented matrix back to the corresponding equations yields

\[
\begin{align*}
x + 3y + 0z &= b_1 - 3b_2 + 0b_3 \\
0x + 0y + z &= 3b_1 - 8b_2 + 0b_3 \\
0x + 0y + 0z &= -b_1 + 6b_2 + b_3
\end{align*}
\]
or

\[
\begin{align*}
x + 3y &= b_1 - 3b_2 \\
z &= 3b_1 - 8b_2 \\
0 &= -b_1 + 6b_2 + b_3.
\end{align*}
\]

We can see that this system does not have any solutions (in \( x, y, \) and \( z \)) if

\[-b_1 + 6b_2 + b_3 \neq 0.\]

Hence, for such a \( \vec{b} \), the system \( A\vec{x} = \vec{b} \) has no solutions, and hence such a \( \vec{b} \) cannot be in the range of the function \( f(\vec{x}) = A\vec{x} \).

On the other hand, if

\[-b_1 + 6b_2 + b_3 = 0. \quad (1)\]

holds then \( \vec{b} \) is such that \( A\vec{x} = \vec{b} \) does have solutions, and so \( \vec{b} \) is in the range of \( f(\vec{x}) = A\vec{x} \).

In the exercises on this topic, let’s refer to this equation (1), and its analogs, as the equation condition(s) on \( \vec{b} \) so that \( A\vec{x} = \vec{b} \) has a solution. We have “condition(s)” here because there could be more than one equation.

Furthermore, we can write down a formula for all the solutions (setting free variable \( y \) equal to \( t \)):

\[
\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_1 - 3b_2 - 3t \\ t \\ 3b_1 - 6b_2 \end{bmatrix}.
\]
For another interpretation of this situation, we recall (from section 3) that the product \( A\vec{x} \) is a linear combination of columns of \( A \) with coefficients furnished by the entries of vector \( \vec{x} \).

Thus the following are the same thing:

(i) \( \vec{b} \) is in the range of \( f(\vec{x}) = A\vec{x} \).

(ii) \( A\vec{x} = \vec{b} \) has a solution.

(iii) \( \vec{b} \) can be expressed as a linear combination of the columns of \( A \).

In our example above, \( \vec{b} \), which does \(-b_1 + 6b_2 + b_3 = 0\), can be written as a linear combination of columns of \( A \) like this:

\[
\vec{b} = (b_1 - 3b_2 - 3t)\vec{A}_1 + (t)\vec{A}_2 + (3b_1 - 6b_2)\vec{A}_3
\]

\[
= (b_1 - 3b_2 - 3t) \begin{bmatrix} -8 \\ -3 \\ 10 \end{bmatrix} + (t) \begin{bmatrix} -24 \\ -9 \\ 30 \end{bmatrix} + (3b_1 - 6b_2) \begin{bmatrix} 3 \\ 1 \\ -3 \end{bmatrix}.
\]

By taking free variable \( t \) as zero, we can express all linear combinations of all three columns of \( A \) as linear combinations of the first and last column of \( A \).

We will return to this situation to squeeze more interpretations out of it later.