3 Matrix Operations

There are several operations one can apply to a matrix. Addition and scalar multiplication behave as you would expect (just like in vectors), but matrix multiplication and its counterpart matrix inverses are more interesting.

3.1 Basic Matrix Operations

We need some notation. For matrix A, the notation a_{ij} means the entry in row i and column j of A. (Always row index first.)

Matrix ***addition*** *requires that the two matrices have the same dimensions. The sum is defined by adding the corresponding entries. For example,*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+a_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+a_{22} & a_{23}+b_{23} \end{bmatrix}$$

Similarly, ***scalar multiplication*** *is defined entry-wise. For example,*

$$c\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} \\ ca_{21} & ca_{22} \end{bmatrix}$$

Example.

$$3\begin{bmatrix} 2 & 2 \\ 1 & 4 \end{bmatrix} + 4\begin{bmatrix} 0 & -1 \\ -5 & 7 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 3 & 12 \end{bmatrix} + \begin{bmatrix} 0 & -4 \\ -20 & 28 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ -17 & 40 \end{bmatrix}$$

Another matrix operation is the transpose:

The ***transpose*** *of a matrix A, denoted A^T, exchanges rows and columns. That is, $(A^T)_{ij} = A_{ji}$.*

Example.

$$\text{The transpose of } \begin{bmatrix} 3 & 4 & 7 \\ -2 & 5 & -3 \end{bmatrix} \text{ is } \begin{bmatrix} 3 & -2 \\ 4 & 5 \\ 7 & -3 \end{bmatrix}$$

A ***square*** *matrix has an equal number of rows and columns. The* ***diagonal*** *of a square matrix runs from top-left to bottom-right. A* ***symmetric matrix*** *is a square matrix that is symmetric around its diagonal. In other words, $A = A^T$.*

3.2 Matrix Multiplication

An important operation is matrix multiplication. It looks intimidating to start with, but you'll get used to it. Matrix **multiplication** produces a matrix. Only matrices of compatible sizes can be multiplied. There are multiple ways to present the definition. One way to define matrix-matrix multiplication is in terms of matrix-vector multiplication:

If matrix A is $m \times n$ and matrix B is $r \times s$, then for the product AB to be valid it must be that $n = r$. If valid, the product AB has size $m \times s$. The columns of the product are the results of multiplying the first matrix by the columns of the second. That is,

$$AB = \begin{bmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \cdots & A\boldsymbol{b}_s \end{bmatrix}$$

where $\boldsymbol{b}_j$ is the j^{th} column of B.

EXAMPLE. *Here is the product of a 2×3 and 3×4 matrix:*

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & -2 \end{bmatrix} \begin{bmatrix} 3 & 1 & -1 & 5 \\ -2 & 0 & 3 & -4 \\ 1 & -2 & 2 & -1 \end{bmatrix} = \begin{bmatrix} -2 & 3 & 3 & -2 \\ -8 & 4 & 5 & -10 \end{bmatrix}$$

An example detail: the 3rd column of the result is given by $-\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 3 \end{bmatrix} + 2\begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

It is more usual to define matrix multiplication without reference to vectors. There is a formula for each entry in the product; namely the sum

$$(AB)_{ij} = \sum_k a_{ik} b_{kj}$$

That is, to calculate the entry in row i and column j of the product AB, look at row i of the first matrix A and column j of the second matrix B; then multiply the corresponding entries and add. This calculation is illustrated here:

Example. *The entry in the 2nd row 3rd column of the previous example is calculated as* $0 \times (-1) + 3 \times 3 + (-2) \times 2 = 5$.

$$\begin{bmatrix} \bullet & \bullet & \bullet \\ 0 & 3 & -2 \end{bmatrix} \begin{bmatrix} \bullet & \bullet & -1 & \bullet \\ \bullet & \bullet & 3 & \bullet \\ \bullet & \bullet & 2 & \bullet \end{bmatrix} = \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & 5 & \bullet \end{bmatrix}$$

$$(0 \times -1) + (3 \times 3) + (-2 \times 2)$$

3.3 Properties of Matrix Multiplication

It is very important to note two fundamental properties about multiplication:

1. Matrix multiplication is **associative**. That is, brackets don't matter. For example, the two products $(AB)C$ and $A(BC)$ are equal (and the one product is valid whenever the other one is).

2. However, matrix multiplication is not **commutative**. That is, order matters. There is no guarantee that $AB = BA$. Indeed, the one product might be valid when the other one is not. Even if A and B are square matrices of the same size, so that both products are defined and the results have the same size, there is no guarantee (and indeed it is unlikely that) the two products are the same.

We repeat this for emphasis:

Fact 3.1 *Matrix multiplication is associative but not commutative.*

In particular, this means that if we multiply by a matrix, we must specify whether we are multiplying it on the left or the right.

*A **diagonal matrix** is a square matrix that has zeros off the diagonal (and might or might not have zeroes on the diagonal). The **identity matrix** I_n is the $n \times n$ diagonal matrix with 1's on the diagonal. (We sometimes write just I.) Its columns are the vectors $\mathbf{e}_i$: these have 0's in every position except for a 1 in the i^{th} position.*

Example. *If A is a square matrix, then $IA = AI = A$, where I is the identity matrix of the same size. (This is why I is called the identity matrix.)*

Using the above formula for the entries of the product, the following fact about transposes can be shown. Note that the order is swapped!

Fact 3.2 $(AB)^T = B^T A^T$

We conclude this section with the following fact.

Fact 3.3 *Each elementary row operation is equivalent to multiplying on the left by a matrix called an* ***elementary matrix****.*

Instead of providing a proof of this fact, we just give an example of each type.

EXAMPLE. *If A is a 2×2 matrix, then:*
left-multiplication by $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ *interchanges the first and second row;*
left-multiplication by $\begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix}$ *divides the second row by 3; and*
left-multiplication by $\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$ *subtracts twice the first row from the second.*

3.4 Matrix Transforms

We can also think of the matrix multiplication $A\boldsymbol{x}$ as transforming the vector $\boldsymbol{x}$.

If A is an $m \times n$ matrix, then the matrix transform $\boldsymbol{x} \mapsto A\boldsymbol{x}$ takes a vector in $\mathbb{R}^n$ and produces a vector (called its ***image****) in $\mathbb{R}^m$. That is, it is a function with domain $\mathbb{R}^n$ and range contained in $\mathbb{R}^m$.*

EXAMPLE. *For example, if A is* $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, *then the transform maps* $(5, 3)$ *to* $(3, 5)$.

Some matrix transforms have physical or geometric meaning:

EXAMPLE. *Examples of transforms include:*

- *projections, such as* $P = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$
- *shears, such as* $S = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$

- *contractions/dilations, such as* $C = \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$
- *rotations, such as* $R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

The reasons for the names can be seen by observing the effect of these transforms on sets of points in the plane.

One useful property when applying matrix transforms is that the composition of transforms (meaning applying one transform after another) is equivalent to matrix multiplication. For example, in $\mathbb{R}^2$ to rotate by θ and then contract by a factor of 3, transform by the matrix product CR, where the matrices C and R are as in the above example.

One can also think of applying the same transform multiple times. For matrix A, we use A^p to mean the product of p copies of A. (This needs A to be square.) That is,

$$A^p = \underbrace{A\,A\,A \cdots A}_{p \text{ copies}}$$

3.5 The Inverse of a Matrix

We next define the inverse of a matrix.

*The **inverse** of a square matrix A, denoted A^{-1}, is the matrix such that $A\,A^{-1} = A^{-1}\,A = I$. The inverse is not guaranteed to exist. If it exists, then A is said to be **invertible**; otherwise A is said to be **not invertible** or **singular**.*

It can be shown that, if the inverse exists, it is unique. If matrix A is invertible, then $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution, namely $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. However, perhaps surprisingly, the inverse is not often calculated, though its existence is crucial.

The inverse of a 2×2 matrix has a formula. Note that the formula also captures when the inverse exists: the matrix is invertible if and only if $ad - bc \neq 0$.

Fact 3.4 *The inverse of a 2×2 matrix is given by:*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

EXAMPLE.

$$C = \begin{bmatrix} 2 & -5 \\ -3 & 9 \end{bmatrix} \quad \textit{has inverse} \quad C^{-1} = \frac{1}{2\times 9 - (-5)\times(-3)} \begin{bmatrix} 9 & 5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 5/3 \\ 1 & 2/3 \end{bmatrix}$$

One general way to find the inverse, albeit slow, is to solve the collection of n vector equations $A\boldsymbol{x} = \boldsymbol{e}_1, \ldots, A\boldsymbol{x} = \boldsymbol{e}_n$ (where the $\boldsymbol{e}_j$ are the columns of I_n as before). These n linear systems can be solved simultaneously by augmenting the matrix with the identity matrix I_n, and bringing the result to reduced row echelon form.

EXAMPLE.

$$C = \begin{bmatrix} 2 & -5 \\ -3 & 9 \end{bmatrix} \quad \textit{is augmented to} \quad \left[\begin{array}{cc|cc} 2 & -5 & 1 & 0 \\ -3 & 9 & 0 & 1 \end{array}\right]$$

$$\textit{This reduces to} \quad \left[\begin{array}{cc|cc} 1 & 0 & 3 & 5/3 \\ 0 & 1 & 1 & 2/3 \end{array}\right] \quad \textit{so that} \quad C^{-1} = \begin{bmatrix} 3 & 5/3 \\ 1 & 2/3 \end{bmatrix}$$

The above discussion shows that a matrix is invertible if and only if it is row equivalent to the identity. Indeed, it can be shown that the inverse is the product of the elementary matrices that reduce A to the identity.

Here are some useful formulas for inverses:

Fact 3.5 *If A and B are square matrices of the same size:*
(a) $(A^{-1})^{-1} = A$
(b) $(AB)^{-1} = B^{-1}A^{-1}$ *(Note the reversal!)*
(c) $(A^T)^{-1} = (A^{-1})^T$.

PROOF. *In each case, to prove that one has the inverse, one can just calculate the product and check that this yields the identity matrix. We do part (b); the other parts are similar. This proof uses the fact that, by the associative law, one can move the brackets around. We have*

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} && \textit{by associative law} \\ &= AIA^{-1} && \textit{since } B \textit{ and } B^{-1} \textit{ multiply to } I \\ &= AA^{-1} \\ &= I && \textit{since } A \textit{ and } A^{-1} \textit{ multiply to } I \end{aligned}$$

Multiplying on the left similarly yields the identity matrix.

3.6 Characterization of Invertible Matrices

Let us collect in one place the big theorem that pulls many things together. (We will define the determinant in the next chapter.)

Fact 3.6 *For an $n \times n$ matrix A, the following conditions are* ***equivalent****:*

- *A is invertible*
- *A has n pivots*
- *A is row equivalent to I_n*
- *$A\boldsymbol{x} = \mathbf{0}$ has a unique solution*
- *the columns of A are linearly independent*
- *the span of the columns of A is all of $\mathbb{R}^n$*
- *the determinant of A is nonzero*

We will use this fact repeatedly.

3.7 Block and Diagonal Matrices

A ***partitioned*** *matrix has the rows and columns partitioned, dividing the matrix up into blocks. A* ***block-diagonal*** *matrix is one where all blocks off the diagonal are zero.*

If they are the correct size, the blocks of a partitioned matrix can be treated as elements for formulas.

EXAMPLE. *If two $n \times n$ matrices are partitioned into $n/2 \times n/2$ blocks and both matrices have a zero block as the bottom left, then*

$$\begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \begin{bmatrix} D & E \\ 0 & F \end{bmatrix} = \begin{bmatrix} AD & AE + BF \\ 0 & CF \end{bmatrix}$$

It can be shown that a block-diagonal matrix is invertible if and only if all the diagonal blocks are invertible. Moreover, its inverse is the block-diagonal matrix with the inverses of the diagonal blocks.

EXAMPLE.

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & -5 \\ 0 & -5 & 9 \end{bmatrix}^{-1} = \begin{bmatrix} 1/6 & 0 & 0 \\ 0 & 9/2 & 5/2 \\ 0 & 5/2 & 3/2 \end{bmatrix}$$

A **lower triangular** matrix is one whose entries above the main diagonal are zero. An **upper triangular** matrix is defined similarly. For example, a diagonal matrix is both lower and upper triangular.

Fact 3.7 *A square triangular matrix is invertible if and only if every entry on the diagonal is nonzero.*

This fact follows from the Characterization of Invertible Matrices (Fact 3.6), since there will be a full set of pivots if and only if every entry on the diagonal is nonzero.

Practice

3.1. Compute XY, $X+Y$, YZ, $Y+Z$, and ZX, when they exist, for the following the matrices:

$$X = \begin{bmatrix} -3 & 1 & 4 \\ 1 & 2 & 0 \end{bmatrix} \qquad Y = \begin{bmatrix} 5 & -4 \\ 3 & -1 \end{bmatrix} \qquad Z = \begin{bmatrix} -2 & 7 \\ 5 & -1 \end{bmatrix}$$

3.2. Assume A, B, and C are 2×2 matrices.

(a) Show that if $AB = AC$ and A is invertible, then $B = C$.

(b) Give an example such that $AB = AC$ but $B \neq C$.

3.3. Calculate the inverses of the following matrices:

$$E = \begin{bmatrix} 2 & -2 \\ 3 & 7 \end{bmatrix} \qquad F = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad G = \begin{bmatrix} 2 & -2 & 0 & 0 \\ 3 & 7 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}$$

Solutions to Practice Exercieses

3.1. XY and $X+Y$ do not exist.

$$YZ = \begin{bmatrix} -30 & 39 \\ -11 & 22 \end{bmatrix} \qquad Y+Z = \begin{bmatrix} 3 & 3 \\ 8 & -2 \end{bmatrix} \qquad ZX = \begin{bmatrix} 13 & 12 & -8 \\ -16 & 3 & 20 \end{bmatrix}$$

3.2. (a) If A^{-1} exists, then we have $A^{-1}(AB) = A^{-1}(AC)$ so that, using the associative law, we have $(A^{-1}A)B = (A^{-1}A)C$ so that $B = C$.

(b) e.g. Take A to be the all-zero matrix.

3.3.

$$E^{-1} = \begin{bmatrix} 7/20 & 1/10 \\ -3/20 & 1/10 \end{bmatrix} \qquad F^{-1} = F \qquad G^{-1} = \begin{bmatrix} E^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{9} \end{bmatrix}$$