A Crash Course in (2 × 2) Matrices

Several weeks worth of matrix algebra in an hour… (Relax, we will only study the simplest case, that of 2 × 2 matrices.)

Review topics:

1. What is a matrix (*pl*. matrices)?

> A matrix is a rectangular array of objects (called *entries*). Those entries are usually numbers, but they can also include functions, vectors, or even other matrices. Each entry's position is addressed by the row and column (in that order) where it is located. For example, a_{52} represents the entry positioned at the 5th row and the 2nd column of the matrix $\boldsymbol{A}$.

2. The size of a matrix

> The size of a matrix is specified by 2 numbers
>
> [*number of rows*] × [*number of columns*].
>
> Therefore, an $m \times n$ matrix is a matrix that contains m rows and n columns. A matrix that has equal number of rows and columns is called a *square matrix*. A square matrix of size $n \times n$ is usually referred to simply as a square matrix of size (or order) n.

Notice that if the number of rows or columns is 1, the result (respectively, a $1 \times n$, or an $m \times 1$ matrix) is just a vector. A $1 \times n$ matrix is called a *row vector*, and an $m \times 1$ matrix is called a *column vector*. Therefore, vectors are really just special types of matrices. Hence, you will probably notice the similarities between many of the matrix operations defined below and vector operations that you might be familiar with.

3. Two special types of matrices

Identity matrices (square matrices only)

The $n \times n$ identity matrix is often denoted by $\boldsymbol{I}_n$.

$$\boldsymbol{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \boldsymbol{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ etc.}$$

Properties (assume $\boldsymbol{A}$ and $\boldsymbol{I}$ are of the same size):

$$\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$$
$$\boldsymbol{I}_n\,\boldsymbol{x} = \boldsymbol{x}, \qquad \boldsymbol{x} = \text{any } n \times 1 \text{ vector}$$

Zero matrices – matrices that contain all-zero entries.

Properties:

$$\boldsymbol{A} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{A} = \boldsymbol{A}$$
$$\boldsymbol{A}\,\boldsymbol{0} = \boldsymbol{0} = \boldsymbol{0}\,\boldsymbol{A}$$

4. Arithmetic operations of matrices

(i) Addition / subtraction

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \pm \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a \pm e & b \pm f \\ c \pm g & d \pm h \end{bmatrix}$$

(ii) Scalar Multiplication

$$k\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ka & kb \\ kc & kd \end{bmatrix}, \qquad \text{for any scalar } k.$$

(iii) Matrix multiplication

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{bmatrix}$$

The matrix multiplication $\boldsymbol{AB} = \boldsymbol{C}$ is defined only if there are as many rows in $\boldsymbol{B}$ as there are columns in $\boldsymbol{A}$. For example, when $\boldsymbol{A}$ is $m \times k$ and $\boldsymbol{B}$ is $k \times n$. The product matrix $\boldsymbol{C}$ is going to be of size $m \times n$, and whose ij-th entry, c_{ij}, is equal to the vector dot product between the i-th row of $\boldsymbol{A}$ and the j-th column of $\boldsymbol{B}$. Since vectors are matrices, we can also multiply together a matrix and a vector, assuming the above restriction on their sizes is met. The product of a 2×2 matrix and a 2-entry column vector is

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax+by \\ cx+dy \end{bmatrix}.$$

Note 1: Two square matrices of the same size can always be multiplied together. Because, obviously, having the same number of rows and columns, they satisfy the size requirement outlined above.

Note 2: In general, $\boldsymbol{AB} \neq \boldsymbol{BA}$. Indeed, depending on the sizes of $\boldsymbol{A}$ and $\boldsymbol{B}$, one product might not even be defined while the other product is.

5. Determinant (square matrices only)

For a 2 × 2 matrix, its determinant is given by the formula

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

Note: The determinant is a function whose domain is the set of all square matrices of a certain size, and whose range is the set of all real (or complex) numbers.

6. Inverse matrix (of a square matrix)

Given an $n \times n$ square matrix $\boldsymbol{A}$, if there exists a matrix $\boldsymbol{B}$ (necessarily of the same size) such that

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}_n,$$

then the matrix $\boldsymbol{B}$ is called the *inverse matrix* of $\boldsymbol{A}$, denoted $\boldsymbol{A}^{-1}$. The inverse matrix, if it exists, is unique for each $\boldsymbol{A}$. A matrix is called *invertible* if it has an inverse matrix.

Theorem: For any 2 × 2 matrix $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$,
its inverse, if exists, is given by

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

Theorem: A square matrix is invertible if and only if its determinant is nonzero.

Examples: Let $\boldsymbol{A} = \begin{bmatrix} 1 & -2 \\ 5 & 2 \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} 2 & -3 \\ -1 & 4 \end{bmatrix}$.

(i) $$2\boldsymbol{A} - \boldsymbol{B} = 2\begin{bmatrix} 1 & -2 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 2 & -3 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 2-2 & -4-(-3) \\ 10-(-1) & 4-4 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & -1 \\ 11 & 0 \end{bmatrix}$$

(ii) $$\boldsymbol{AB} = \begin{bmatrix} 1 & -2 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 2 & -3 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 2+2 & -3-8 \\ 10-2 & -15+8 \end{bmatrix} = \begin{bmatrix} 4 & -11 \\ 8 & -7 \end{bmatrix}$$

On the other hand:
$$\boldsymbol{BA} = \begin{bmatrix} 2 & -3 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} 2-15 & -4-6 \\ -1+20 & 2+8 \end{bmatrix} = \begin{bmatrix} -13 & -10 \\ 19 & 10 \end{bmatrix}$$

(iii) $\det(\boldsymbol{A}) = 2 - (-10) = 12$, $\det(\boldsymbol{B}) = 8 - 3 = 5$.

Since neither is zero, as a result, they are both invertible matrices.

(iv) $$\boldsymbol{A}^{-1} = \frac{1}{2-(-10)}\begin{bmatrix} 2 & 2 \\ -5 & 1 \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 2 & 2 \\ -5 & 1 \end{bmatrix} = \begin{bmatrix} 1/6 & 1/6 \\ -5/12 & 1/12 \end{bmatrix}$$

7. Systems of linear equations (also known as *linear systems*)

A system of linear (algebraic) equations, $\boldsymbol{Ax} = \boldsymbol{b}$, could have zero, exactly one, or infinitely many solutions. (Recall that each linear equation has a line as its graph. A solution of a linear system is a common intersection point of all the equations' graphs − and there are only 3 ways a set of lines could intersect.)

If the vector $\boldsymbol{b}$ on the right-hand side is the zero vector, then the system is called homogeneous. A homogeneous linear system always has a solution, namely the all-zero solution (that is, the origin). This solution is called the *trivial solution* of the system. Therefore, a homogeneous linear system $\boldsymbol{Ax} = \boldsymbol{0}$ could have either exactly one solution, or infinitely many solutions. There is no other possibility, since it always has, at least, the trivial solution. If such a system has n equations and exactly the same number of unknowns, then the number of solution(s) the system has can be determined, without having to solve the system, by the determinant of its coefficient matrix:

> *Theorem*: If $\boldsymbol{A}$ is an $n \times n$ matrix, then the homogeneous linear system $\boldsymbol{Ax} = \boldsymbol{0}$ has exactly one solution (the trivial solution) if and only if $\boldsymbol{A}$ is invertible (that is, it has a nonzero determinant). It will have infinitely many solutions (the trivial solution, plus infinitely many nonzero solutions) if $\boldsymbol{A}$ is not invertible (equivalently, has zero determinant).

8. Eigenvalues and Eigenvectors

Given a square matrix $\boldsymbol{A}$, suppose there are a constant r and a nonzero vector $\boldsymbol{x}$ such that

$$\boldsymbol{Ax} = r\boldsymbol{x},$$

then r is called an *Eigenvalue* of $\boldsymbol{A}$, and $\boldsymbol{x}$ is an *Eigenvector* of $\boldsymbol{A}$ corresponding to r.

Do eigenvalues/vectors always exist for any given square matrix? The answer is yes. How do we find them, then?

Rewrite the above equation, we get $\boldsymbol{Ax} - r\boldsymbol{x} = \boldsymbol{0}$. The next step would be to factor out $\boldsymbol{x}$. But doing so would give the expression

$$(\boldsymbol{A} - r)\boldsymbol{x} = \boldsymbol{0}.$$

Notice that it requires us to subtract a number from an $n \times n$ matrix. That's an undefined operation. Hence, we need to further refined it by rewriting the term $r\boldsymbol{x} = r\boldsymbol{I}\boldsymbol{x}$, and then factoring out $\boldsymbol{x}$, obtaining

$$(\boldsymbol{A} - r\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}.$$

This is an $n \times n$ system of homogeneous linear (algebraic) equations, where the coefficient matrix is $(\boldsymbol{A} - r\boldsymbol{I})$. We are looking for a nonzero solution $\boldsymbol{x}$ of this system. Hence, by the theorem we have just seen, the necessary and sufficient condition for the existence of such a nonzero solution, which will become an eigenvector of $\boldsymbol{A}$, is that the coefficient matrix $(\boldsymbol{A} - r\boldsymbol{I})$ must have zero determinant. Set its determinant to zero and what we get is a degree n polynomial equation in terms of r. The case of a 2×2 matrix is as follow:

$$\boldsymbol{A} - r\boldsymbol{I} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} - r\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a - r & b \\ c & d - r \end{bmatrix}.$$

Its determinant, set to 0, yields the equation

$$\det\begin{bmatrix} a-r & b \\ c & d-r \end{bmatrix} = (a-r)(d-r) - bc = r^2 - (a+d)r + (ad-bc) = 0$$

It is a degree 2 polynomial equation of r, as you can see.

This polynomial on the left is called the *characteristic polynomial* of the (original) matrix $\boldsymbol{A}$, and the equation is the *characteristic equation* of $\boldsymbol{A}$. The root(s) of the characteristic polynomial are the eigenvalues of $\boldsymbol{A}$. Since any degree n polynomial always has n roots (real and/or complex; not necessarily distinct), any $n \times n$ matrix always has at least one, and up to n different eigenvalues.

Once we have found the eigenvalue(s) of the given matrix, we put each specific eigenvalue back into the linear system $(\boldsymbol{A} - r\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$ to find the corresponding eigenvectors.

Examples: $\boldsymbol{A} = \begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix}$

$$\boldsymbol{A} - r\boldsymbol{I} = \begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix} - r\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2-r & 3 \\ 4 & 3-r \end{bmatrix}.$$

Its characteristic equation is

$$\det\begin{bmatrix} 2-r & 3 \\ 4 & 3-r \end{bmatrix} = (2-r)(3-r) - 12 = r^2 - 5r - 6 = (r+1)(r-6) = 0$$

The eigenvalues are, therefore, $r = -1$ and 6.

Next, we will substitute each of the 2 eigenvalues into the matrix equation $(\boldsymbol{A} - r\boldsymbol{I})\,\boldsymbol{x} = \boldsymbol{0}$.

For $r = -1$, the system of linear equations is

$$(\boldsymbol{A} - r\boldsymbol{I})\,\boldsymbol{x} = (\boldsymbol{A} + \boldsymbol{I})\,\boldsymbol{x} = \begin{bmatrix} 2+1 & 3 \\ 4 & 3+1 \end{bmatrix} x = \begin{bmatrix} 3 & 3 \\ 4 & 4 \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Notice that the matrix equation represents a degenerated system of 2 linear equations. Both equations are constant multiples of the equation $x_1 + x_2 = 0$. There is now only 1 equation for the 2 unknowns, therefore, there are infinitely many possible solutions. <u>This is always the case when solving for eigenvectors</u>. Necessarily, there are infinitely many eigenvectors corresponding to each eigenvalue.

Solving the equation $x_1 + x_2 = 0$, we get the relation $x_2 = -x_1$. Hence, the eigenvectors corresponding to $r = -1$ are all nonzero multiples of

$$k_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Similarly, for $r = 6$, the system of equations is

$$(\boldsymbol{A} - r\boldsymbol{I})\,\boldsymbol{x} = (\boldsymbol{A} - 6\boldsymbol{I})\,\boldsymbol{x} = \begin{bmatrix} 2-6 & 3 \\ 4 & 3-6 \end{bmatrix} x = \begin{bmatrix} -4 & 3 \\ 4 & -3 \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Both equations in this second linear system are equivalent to $4x_1 - 3x_2 = 0$. Its solutions are given by the relation $4x_1 = 3x_2$. Hence, the eigenvectors corresponding to $r = 6$ are all nonzero multiples of

$$k_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

Note: Every nonzero multiple of an eigenvector is also an eigenvector.

Two short-cuts to find eigenvalues:

1. If $\boldsymbol{A}$ is a diagonal or triangular matrix, that is, if it has the form

$$\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}, \text{ or } \quad \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}, \text{ or } \quad \begin{bmatrix} a & 0 \\ c & d \end{bmatrix}.$$

Then the eigenvalues are just the main diagonal entries, $r = a$ and d in all 3 examples above.

2. If $\boldsymbol{A}$ is any 2×2 matrix, then its characteristic equation is

$$\det \begin{bmatrix} a-r & b \\ c & d-r \end{bmatrix} = r^2 - (a+d)r + (ad-bc) = 0$$

If you are familiar with terminology of linear algebra, the characteristic equation can be memorized rather easily as

$$\boxed{r^2 - \text{Trace}(\boldsymbol{A})\, r + \det(\boldsymbol{A}) = 0.}$$

Note: For any square matrix $\boldsymbol{A}$, Trace($\boldsymbol{A}$) = [*sum of all entries on the main diagonal (running from top-left to bottom-right)*]. For a 2×2 matrix $\boldsymbol{A}$, Trace($\boldsymbol{A}$) = $a + d$.

A short-cut to find eigenvectors (of a 2×2 matrix):

Similarly, there is a trick that enables us to find the eigenvectors of any 2×2 matrix without having to go through the whole process of solving systems of linear equations. This short-cut is especially handy when the eigenvalues are complex numbers, since it avoids the need to solve the linear equations which will have complex number coefficients. (*Warning*: *This method does not work for any matrix of size larger than* 2×2.)

We first find the eigenvalue(s) and then write down, for each eigenvalue, the matrix $(\boldsymbol{A} - r\boldsymbol{I})$ as usual. Then we take *any* row of $(\boldsymbol{A} - r\boldsymbol{I})$ that is not consisted of entirely zero entries, say it is the row vector (α, β). We put a minus sign in front of one of the entries, for example, $(\alpha, -\beta)$. Then an eigenvector of the matrix A is found by switching the two entries in the above vector, that is, $\boldsymbol{k} = (-\beta, \alpha)$.

Example: Previously, we have seen $\boldsymbol{A} = \begin{bmatrix} 2 & 3 \\ 4 & 3 \end{bmatrix}$.

The characteristic equation is

$$r^2 - \mathrm{Trace}(\boldsymbol{A})\, r + \det(\boldsymbol{A}) = r^2 - 5r - 6 = (r+1)(r-6) = 0,$$

which has roots $r = -1$ and 6. For $r = -1$, the matrix $(\boldsymbol{A} - r\boldsymbol{I})$ is $\begin{bmatrix} 3 & 3 \\ 4 & 4 \end{bmatrix}$.

Take the first row, (3, 3), which is a non-zero vector; put a minus sign to the first entry to get $(-3, 3)$; then switch the entry, we now have $\boldsymbol{k}_1 = (3, -3)$. It is indeed an eigenvector, since it is a nonzero constant multiple of the vector we found earlier.

On very rare occasions, both rows of the matrix $(\boldsymbol{A} - r\boldsymbol{I})$ have all zero entries. If so, the above algorithm will not be able to find an eigenvector. Instead, under this circumstance any non-zero vector will be an eigenvector.

