Matrices

Joshua Ballew

Spring 2016

Matrices are arrays of numbers. More specifically, an $m \times n$ matrix A is an array of m rows and n columns. We denote the element in the i-th row and j-th column as a_{ij} or $(A)_{ij}$, calling it the ij-th element of A. We define the transpose of A, denoted A^T (your book uses A') as the $n \times m$ matrix with $(A^T)_{ij} = a_{ji}$.

1 Scalar Multiplication and Addition

For any $m \times n$ matrix A and constant c, we define the scalar product of cA as the $m \times n$ matrix where $(cA)_{ij} = ca_{ij}$. We define $-A$ as $(-1)A$, the matrix formed by multiplying A by the scalar -1.

Two matrices can be added together provided they are the same size. For two $m \times n$ matrices A and B, we define the sum $A+B$ such that $(A+B)_{ij} = a_{ij}+b_{ij}$. The difference $A - B$ is defined as $A + (-1)B$.

Scalar multiplication and matrix addition obey the following properties.

Theorem 1. *Let A, B, and C be $m \times n$ matrices, O be the $m \times n$ matrix with all entries being zero, and let c and d be real numbers. Then*

1. $(A + B) + C = A + (B + C)$
2. $A + B = B + A$
3. $A + O = A$
4. $A - A = O$
5. $(c + d)A = cA + dA$
6. $c(A + B) = cA + cB$.

2 Matrix Multiplication

In addition to scalar multiplication and addition of matrices, we can also multiply two matrices. However, we can only multiply two matrices subject to conditions on their sizes. The product of two matrices A and B, denoted AB, is defined only if A is $m \times r$ and B is $r \times n$. That is, the number of columns in the first matrix must be equal to the number of rows in the second matrix. The resulting matrix AB will be $m \times n$. We define the product AB such that

$$(AB)_{ij} = \sum_{k=1}^{r} a_{ik} b_{kj}.$$

In other words, for the ij-th element of the product, we take the first entry in the i-th row of A and multiply it by the first element in the j-th column of B, the second entry in the i-th row of A multiplied it by the second element in the j-th column of B, and so on, and then add these r products together.

Matrix multiplication obeys the following properties.

Theorem 2. *Let A be $m \times r$, B and D be $r \times s$, and C be $s \times n$. Let I_k be the $k \times k$ square matrix with ones on the main diagonal (upper left to lower right) and zeros for all the other entries. Then*

1. *$(AB)C = A(BC)$, which is often denoted ABC (what are the dimensions of ABC?)*
2. *AB **need not be equal to** BA (in fact, BA may not be defined)*
3. *$A(B + D) = AB + AD$*
4. *$(B + D)C = BC + DC$*
5. *$AI_r = A$*
6. *$I_m A = A$.*

It is also important to note that if we define

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ am1 & a_{m2} & \dots & a_{mn} \end{bmatrix},$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

and

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

then $A\mathbf{x} = \mathbf{b}$ is the same thing as the system of equations

$$\begin{array}{lllll} a_{11}x_1 & +a_{12}x_2 & +\dots & +a_{1n}x_n & = b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\dots & +a_{1n}x_n & = b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & +\dots & +a_{mn}x_n & = b_m. \end{array}$$

3 Vectors

Vectors are matrices of one column, that is, $n \times 1$ matrices. We say that a vector $\mathbf{x} \in \mathbb{R}^n$ if it is an $n \times 1$ matrix. We write the elements of $\mathbf{x}$ as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

We have the above results on matrix addition, scalar multiplication, and matrix multiplication since vectors are matrices. However, we can define a new operation: the dot product. If $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$, we define the dot product as

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

Using the dot product, we define the length of a vector $\mathbf{x}$ as

$$|\mathbf{x}| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}.$$

We also define the angle θ between two non-zero vectors as

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}||\mathbf{y}|}.$$

4 Square Matrices and Eigenvalues and Eigenvectors

For a square $(n \times n)$ matrix A, we define the trace of A, denoted $\operatorname{tr} A$ as

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii},$$

which is just the sum of the elements on the main diagonal.

The determinant of a 2×2 matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is $ad - bc$.

An eigenvalue of a square matrix A is defined as a value λ such that there is some non-zero $\mathbf{x} \in \mathbb{R}^n$ such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Zero can be an eigenvalue, as can complex numbers. Only square matrices can have eigenvalues. To find the eigenvalues of an $n \times n$ matrix A, we solve the equation

$$\det(\lambda I_n - A) = 0.$$

Once we find the eigenvalue(s), we can find corresponding eigenvectors by solving the system of equations

$$\begin{aligned} (\lambda - a)x_1 - bx_2 &= 0 \\ -cx_1 + (\lambda - d)x_2 &= 0 \end{aligned}$$

for each eigenvector λ. Note that when you solve this system, you will always get a free variable. If you don't, there is a mistake in the calculations.

Example 1. Find the eigenvalues and corresponding eigenvectors for the matrix

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}.$$

Solution. We begin by finding $\lambda I_2 - A$ to be

$$\begin{bmatrix} \lambda - 1 & 1 \\ -2 & \lambda - 4 \end{bmatrix}.$$

Its determinant is

$$(\lambda - 1)(\lambda - 4) + 2 = \lambda^2 - 5\lambda + 6.$$

Setting this determinant equal to zero and solving for λ yields solutions $\lambda_1 = 3$ and $\lambda_2 = 2$. To find an eigenvector $\mathbf{u}_1$ corresponding to λ_1, we solve the system

$$\begin{aligned} 2x_1 + x_2 &= 0 \\ -2x_1 - x_2 &= 0, \end{aligned}$$

yielding solutions of the form

$$t\begin{bmatrix} 1 \\ -2 \end{bmatrix},$$

from which we take $t = 1$ to pick $\mathbf{u}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$. (Note: We could pick any non-zero value for t to obtain $\mathbf{u}_1$.) By following a similar procedure for finding an eigenvector $\mathbf{u}_2$ corresponding to λ_2, we find such a vector to be $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ (try this for yourself). $\square$