8. Matrices

Definition An $r \times k$ matrix $M = (m^i_j)$ for $i = 1, \ldots, r; j = 1, \ldots, k$ is a rectangular array of real (or complex) numbers:

$$M = \begin{pmatrix} m^1_1 & m^1_2 & \cdots & m^1_k \\ m^2_1 & m^2_2 & \cdots & m^2_k \\ \vdots & \vdots & & \vdots \\ m^r_1 & m^r_2 & \cdots & m^r_k \end{pmatrix}$$

The numbers m^i_j are called *entries*. The superscript indexes the row of the matrix and the subscript indexes the column of the matrix in which m^i_j appears.

It is often useful to consider matrices whose entries are more general than the real numbers, so we allow that possibility.

An $r \times 1$ matrix $v = (v^r_1) = (v^r)$ is called a *column vector*, written $v = \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^r \end{pmatrix}$. A $1 \times k$ matrix is (perhaps unsurprisingly) called a *row vector*.

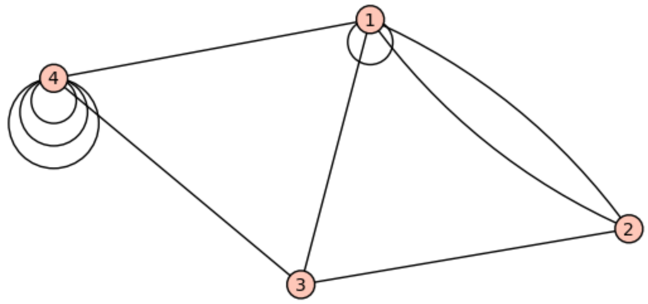
Matrices are a useful way to store information.

Example In computer graphics, you may have encountered image files with a *.gif* extension. These files are actually just matrices: at the start of the file the size of the matrix is given, and then each entry of the matrix is a number indicating the color of a particular pixel in the image.

The resulting matrix is then has its rows shuffled a bit: by listing, say, every eighth row, then a web browser downloading the file can start displaying an incomplete version of the picture before the download is complete.

Finally, a compression algorithm is applied to the matrix to reduce the size of the file.

Example Graphs occur in many applications, ranging from telephone networks to airline routes. In the subject of *graph theory*, a graph is just a collection of vertices and some edges connecting vertices. A matrix can be used to indicate how many edges attach one vertex to another.



For example, the graph pictured above would have the following matrix, where m^i_j indicates the number of edges between the vertices labeled i and j:

$$M = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 3 \end{pmatrix}$$

This is an example of a *symmetric matrix*, since $m^i_j = m^j_i$.

The space of $r \times k$ matrices M^r_k is a vector space with the addition and scalar multiplication defined as follows:

$$M + N = (m^i_j) + (n^i_j) = (m^i_j + n^i_j)$$
$$rM = r(m^i_j) = (rm^i_j)$$

In other words, addition just adds corresponding entries in two matrices, and scalar multiplication multiplies every entry.

Notice that $M^n_1 = \mathbb{R}^n$ is just the vector space of column vectors.

Recall that $r \times k$ matrices can be used for linear transformations $\mathbb{R}^k \to \mathbb{R}^r$ via $MV = (\sum_{j=1}^k m^i_j v_j)$. Here we multiply an $r \times k$ matrix by a $k \times 1$ vector.

Likewise, we can use matrices to represent linear transformations

$$M^s_k \xrightarrow{N} M^r_k$$

via $(LM)^i_l = (\sum_{j=1}^k n^i_j m^j_l)$. This rule obeys linearity.

Notice that in order for the multiplication to make sense, the columns and rows must match. For an $r \times k$ matrix M and an $s \times m$ matrix N, then to make the product MN we must have $k = s$. Likewise, for the product

NM, it is required that $m = r$. A common shorthand for keeping track of the sizes of the matrices involved in a given product is:

$$(r \times k) \times (k \times m) = (r \times m)$$

Example Multiplying a (3×1) matrix and a (1×2) matrix yields a (3×2) matrix.

$$\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 & 1 \cdot 3 \\ 3 \cdot 2 & 3 \cdot 3 \\ 2 \cdot 2 & 2 \cdot 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 6 & 9 \\ 4 & 6 \end{pmatrix}$$

Matrix Terminology The entries m^i_i are called *diagonal*, and the set $\{m^1_1, m^2_2, \ldots\}$ is called the *diagonal of the matrix*.

Any $r \times r$ matrix is called a *square matrix*. A square matrix that is zero for all non-diagonal entries is called a diagonal matrix.

The $r \times r$ diagonal matrix with all diagonal entries equal to 1 is called the *identity matrix*, I_r, or just $\mathbb{1}$. The identity matrix is spacial because $I_r M = M I_k = M$ for all M of size $r \times k$.

In the product matrix in the example above, the diagonal entries are 2 and 9.

An example of a diagonal matrix is $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

Definition The *transpose* of an $r \times k$ matrix $M = (m^i_j)$ is the $k \times r$ matrix with entries

$$M^T = (\bar{m}^i_j)$$

with $\bar{m}^i_j = m^j_i$.

A matrix M is *symmetric* if $M = M^T$.

Example $\begin{pmatrix} 2 & 5 & 6 \\ 1 & 3 & 4 \end{pmatrix}^T = \begin{pmatrix} 2 & 1 \\ 5 & 3 \\ 6 & 4 \end{pmatrix}$

Observations

- Only square matrices can be symmetric.
- The transpose of a column vector is a row vector, and vice-versa.
- Taking the transpose of a matrix twice does nothing. *i.e.,* $(M^T)^T = M$.

Theorem 0.1 (Transpose and Multiplication)**.** *Let M, N be matrices such that MN makes sense. Then $(MN)^T = N^T M^T$.*

References

Hefferon, Chapter Three, Section IV, parts 1-3.
Wikipedia:

- Matrix Multiplication

Review Questions

1. Above, we showed that *left* multiplication by an $r \times k$ matrix N was a linear transformation $M^s_k \xrightarrow{N} M^r_k$. Show that *right* multiplication by an $s \times m$ matrix R is a linear transformation $M^s_k \xrightarrow{R} M^s_m$. In other words, show that right matrix multiplication obeys linearity.

2. Prove the theorem $(MN)^T = N^T M^T$.

3. Let $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ be column vectors. Prove that the dot product $x \cdot y = x^T \mathbb{1} y$.

4. Explain what happens to a matrix when:

   (i) You multiply it on the left by a diagonal matrix.

   (ii) You multiply it on the right by a diagonal matrix.

   Give a few simple examples before you start explaining.