

Chapter 19. General Matrices

An $n \times m$ matrix is an array

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} = [a_{ij}].$$

The matrix A has n **row vectors**

$$\text{row}_i(A) = [a_{i1}, a_{i2}, \dots, a_{im}] \in \mathbb{R}^m$$

and m **column vectors**

$$\text{col}_j(A) = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{bmatrix} \in \mathbb{R}^n.$$

The ij entry a_{ij} is the number in row i , column j .

If $A = [a_{ij}]$ is $n \times m$ and $B = [b_{ij}]$ is $m \times k$, then the product matrix AB is defined by the formula

$$ij \text{ entry of } AB = \sum_{\ell=1}^m a_{i\ell}b_{\ell j} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} = \langle \text{row}_i(A), \text{col}_j(B) \rangle.$$

Note that the number m of columns of A must equal the number of rows of B , otherwise, the product AB is not defined. The inner dimensions must match. The product then has the outer dimensions.

$$\text{Mnemonic : } (n \times m)(m \times k) = n \times k.$$

For example, we can multiply our $n \times m$ matrix A times an $m \times 1$ matrix (which is a vector in \mathbb{R}^m), and we get an $n \times 1$ matrix (which is a vector in \mathbb{R}^n). In other words, A is a function

$$A : \mathbb{R}^m \longrightarrow \mathbb{R}^n$$

which has the **linearity property**: $A(s\mathbf{u} + t\mathbf{v}) = sA\mathbf{u} + tA\mathbf{v}$, for all \mathbf{u}, \mathbf{v} in \mathbb{R}^m and s, t in \mathbb{R} . Any linear function from \mathbb{R}^m to \mathbb{R}^n is multiplication by some matrix A .

If $\mathbf{u}_j = \text{col}_j(A)$, and $\mathbf{x} = (x_1, \dots, x_m)$, then

$$A\mathbf{x} = x_1\mathbf{u}_1 + \dots + x_m\mathbf{u}_m \in \mathbb{R}^n.$$

In particular, if \mathbf{e}_j is the vector with 1 in the j^{th} spot and zeros in all the other spots, then we have the familiar fact that the columns of A are given by

$$A\mathbf{e}_j = \text{col}_j(A). \tag{1}$$

As before, the **kernel** of A is

$$\ker A = \{\mathbf{x} \in \mathbb{R}^m : A\mathbf{x} = \mathbf{0}\} \subset \mathbb{R}^m.$$

If $n = 1$, then $A = [a_1 \ a_2 \ \dots \ a_m]$ and $\ker A$ is the **hyperplane**

$$a_1x_1 + a_2x_2 + \dots + a_mx_m = 0.$$

The kernel of a general $n \times m$ matrix A is the intersection of the row hyperplanes

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{im}x_m = 0, \quad i = 1, 2, \dots, n.$$

The **image** of A is

$$\text{im } A = \{A\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \in \mathbb{R}^m\} \subset \mathbb{R}^n.$$

That is, $\text{im } A$ consists of the vectors in \mathbb{R}^n that come from \mathbb{R}^m via A . From equation (1), it follows that $\text{im } A$ is the span of the columns of A . The kernel and image are related to each other by the

Kernel-Image Theorem: *For any $n \times m$ matrix A , we have*

$$\dim \ker A + \dim \text{im } A = m.$$

For small n, m we have an *ad hoc* understanding of “dimension”. The general definition of dimension will be given in the next chapter.

Intuitively, you can think of A as a messenger, carrying information. In this imaginary viewpoint, the kernel of A is the information lost by A and $\text{im } A$ is the information retained by A . The starting-place \mathbb{R}^m is the total information. Then the Kernel-Image Theorem says:

$$\text{Information lost} + \text{information retained} = \text{total information}.$$

Example 1:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

Here $A : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, sending $\mathbf{e}_1 \rightarrow \mathbf{u} = (1, 3, 5)$, $\mathbf{e}_2 \rightarrow \mathbf{v} = (2, 4, 6)$, and A sends a general vector $(x, y) \in \mathbb{R}^2$ to the linear combination

$$(x, y) \rightarrow x_1\mathbf{u} + x_2\mathbf{v}.$$

Thus, A is an embedding of a plane in \mathbb{R}^3 , namely the plane spanned by \mathbf{u}, \mathbf{v} . Recall that we can find the equation of this plane, take the cross product

$$\mathbf{u} \times \mathbf{v} = (-2, 4, -2),$$

which we can scale to $(1, -2, 1)$. Thus the equation of the plane is

$$x - 2y + z = 0.$$

This plane is called the **image** of A .

In general, the image of a matrix $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is the set of all vectors in \mathbb{R}^n of the form $A\mathbf{u}$, for some $\mathbf{u} \in \mathbb{R}^m$.

Example 2:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix}$$

Here we don't get a plane because the columns are proportional. Instead, every vector of the form $A\mathbf{u}$ is a scalar multiple of $\mathbf{u} = (1, 2, 3)$, so the image is the line $\mathbb{R}\mathbf{u}$. In this case, A is squashing \mathbb{R}^2 into a line in \mathbb{R}^3 .

Example 3:

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

sends $\mathbf{e}_1 \rightarrow \mathbf{e}_1$, $\mathbf{e}_2 \rightarrow \mathbf{e}_2$, $\mathbf{e}_3 \rightarrow \mathbf{0}$. An arbitrary vector $\mathbf{u} = (x, y, z) \in \mathbb{R}^3$ is sent by A to $A\mathbf{u} = (x, y)$. Every vector in \mathbb{R}^3 is squashed vertically onto the floor, which is \mathbb{R}^2 . The image of A is the whole \mathbb{R}^2 . In particular, the cube whose vertices are the vectors with entries 0,1 is squashed straight down into the square in \mathbb{R}^2 whose vertices have entries 0,1.

Example 4:

$$A = \begin{bmatrix} 1 & 0 & x \\ 0 & 1 & y \end{bmatrix}$$

For some choice of (x, y) . This sends the 0,1 cube to an image of the cube in \mathbb{R}^2 that can be seen by Miserable 2-Dimensional Creatures. For example, A sends $(1, 0, 1)$ to the vector $(1, 0) + (x, y)$, and the edge from $(1, 0, 0)$ to $(1, 0, 1)$ to the line segment from $(1, 0)$ to $(1, 0) + (x, y)$. As the M2DC varies (x, y) , it gets different views of the cube, but with the angles deformed. The same thing happens when we try to draw a cube on paper. The price of cramming a 3d object into 2d is loss of information, in this case the angles and relative lengths of the sides. But we can at least see the relation between the vertices and edges and faces.

Exercise 19.1 Compute the product, if can be computed

$$\begin{array}{ll} \text{(a)} \begin{bmatrix} 1 & 0 & 2 \\ 2 & 3 & -1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} & \text{(b)} \begin{bmatrix} 1 & 0 & 2 \\ 2 & 3 & -1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \\ \text{(c)} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} & \end{array}$$

Exercise 19.2 Let $A = [a_{ij}]$ be an arbitrary $n \times n$ matrix, and let D be a diagonal matrix

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

- What are the ij entries of DA and AD ?
- Suppose the d_i 's are distinct, and that $AD = DA$. What can you say about A ?

Exercise 19.3 M2DCs, hence M3DCs like us, can see images of the hypercube in \mathbb{R}^4 , by using a matrix $A : \mathbb{R}^4 \rightarrow \mathbb{R}^2$. First of all, the hypercube has 16 corners (x_1, x_2, x_3, x_4) , $x_i = 0$ or 1. Two corners have an edge between them if they differ in only one component. For example, there is an edge between (1011) and (1001) , but no edge between (1011) and (0111) .

- Choose two random vectors $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ in \mathbb{R}^2 , and let

$$A = \begin{bmatrix} 1 & 0 & u_1 & v_1 \\ 0 & 1 & u_2 & v_2 \end{bmatrix}.$$

(To get a good view, make sure no u_i or v_i is 0 or 1.)

b) Plot the image under A of the 16 corners of the hypercube. You will plot 16 points in the plane \mathbb{R}^2 , but label them by the coordinates of the corresponding points in \mathbb{R}^4 , to keep track of them. For example, the corner (0101) goes via A to $(0, 1) + (v_1, v_2) = (v_1, 1 + v_2)$. Label this point by (0101) .

c) Connect pairs of your 16 points in \mathbb{R}^2 by line segments if their 4-tuple labels differ in only one component.

d) Behold the hypercube!