Chapter Six

Linear Functions and Matrices

6.1 Matrices

Suppose $f:\boldsymbol{R^n} \to \boldsymbol{R^p}$ be a linear function. Let $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ be the coordinate vectors for $\boldsymbol{R^n}$. For any $\boldsymbol{x}$ $\quad \boldsymbol{R^n}$, we have $\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_n\boldsymbol{e}_n$. Thus

$$f(\boldsymbol{x}) = f(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \ldots + x_n\boldsymbol{e}_n) = x_1 f(\boldsymbol{e}_1) + x_2 f(\boldsymbol{e}_2) + \ldots + x_n f(\boldsymbol{e}_n).$$

Meditate on this; it says that a linear function is entirely determined by its values $f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), \ldots, f(\boldsymbol{e}_n)$. Specifically, suppose

$$\begin{aligned} &f(\boldsymbol{e}_1) = (a_{11}, a_{21}, \ldots, a_{p1}), \\ &f(\boldsymbol{e}_2) = (a_{12}, a_{22}, \ldots, a_{p2}), \\ &\vdots \\ &f(\boldsymbol{e}_n) = (a_{1n}, a_{2n}, \ldots, a_{pn}). \end{aligned}$$

Then

$$\begin{aligned} f(\boldsymbol{x}) = (&a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n,\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n, \ldots, \\ &a_{p1}x_1 + a_{p2}x_2 + \ldots + a_{pn}x_n). \end{aligned}$$

The numbers a_{ij} thus tell us everything about the linear function *f*. . To avoid labeling these numbers, we arrange them in a rectangular array, called a ***matrix***:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

The matrix is said to ***represent*** the linear function *f*.

For example, suppose $f : \boldsymbol{R^2} \to \boldsymbol{R^3}$ is given by the receipt

$$f(x_1, x_2) = (2x_1 - x_2,\ x_1 + 5x_2,\ 3x_1 - 2x_2).$$

Then $f(\boldsymbol{e}_1) = f(1,0) = (2,1,3)$, and $f(\boldsymbol{e}_2) = f(0,1) = (-1,5,-2)$. The matrix representing *f* is thus

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

Given the matrix of a linear function, we can use the matrix to compute $f(\boldsymbol{x})$ for any $\boldsymbol{x}$. This calculation is systematized by introducing an arithmetic of matrices. First, we need some jargon. For the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix},$$

the matrices $\left[a_{i1}, a_{i2}, \ldots, a_{in}\right]$ are called ***rows*** of *A*, and the matrices $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{pj} \end{bmatrix}$ are called ***columns*** of *A*. Thus *A* has *p* rows and *n* columns; the ***size*** of *A* is said to be $\boldsymbol{p} \times \boldsymbol{n}$. A vector in $\boldsymbol{R}^{\boldsymbol{n}}$ can be displayed as a matrix in the obvious way, either as a $1 \times n$ matrix, in which case the matrix is called a ***row vector***, or as a $\boldsymbol{n} \times 1$ matrix, called a ***column vector***. Thus the matrix representation of *f* is simply the matrix whose columns are the column vectors $f(\boldsymbol{e}_1), f(\boldsymbol{e}_1), \ldots, f(\boldsymbol{e}_n)$.

Example

Suppose $f: \boldsymbol{R}^{\boldsymbol{3}} \rightarrow \boldsymbol{R}^{\boldsymbol{2}}$ is defined by

$$f(x_1, x_2, x_3) = (2x_1 - 3x_2 + x_3, \ -x_1 + 2x_2 - 5x_3).$$

So $f(\boldsymbol{e}_1) = f(1,0,0) = (2,-1)$, $f(\boldsymbol{e}_2) = f(0,1,0) = (-3,2)$, and $f(\boldsymbol{e}_3) = f(0,0,1) = (1,-5)$. The matrix which represents *f* is thus

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

Now the recipe for computing *f(x)* can be systematized by defining the product of a matrix $\boldsymbol{A}$ and a column vector $\boldsymbol{x}$. Suppose $\boldsymbol{A}$ is a $p \times n$ matrix and $\boldsymbol{x}$ is a $n \times 1$ column vector. For each $i = 1, 2, \ldots, p$, let $\boldsymbol{r}_i$ denote the i^{th} row of $\boldsymbol{A}$. We define the product $\boldsymbol{Ax}$ to be the $p \times 1$ column vector given by

$$\boldsymbol{Ax} = \begin{bmatrix} \boldsymbol{r}_1 \ \boldsymbol{x} \\ \boldsymbol{r}_2 \ \boldsymbol{x} \\ \vdots \\ \boldsymbol{r}_p \ \boldsymbol{x} \end{bmatrix}.$$

If we consider all vectors to be represented by column vectors, then a linear function $f : \boldsymbol{R}^n \to \boldsymbol{R}^p$ is given by $f(\boldsymbol{x}) = \boldsymbol{Ax}$, where, of course, $\boldsymbol{A}$ is the matrix representation of f.

Example

Consider the preceding example:

$$f(x_1, x_2, x_3) = (2x_1 - 3x_2 + x_3, \ -x_1 + 2x_2 - 5x_3).$$

We found the matrix representing f to be

$$\boldsymbol{A} = \begin{bmatrix} 2 & -3 & 1 \\ -1 & 2 & -5 \end{bmatrix}.$$

Then

$$\boldsymbol{Ax} = \begin{bmatrix} 2 & -3 & 1 \\ -1 & 2 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 - 3x_2 + x_3 \\ -x_1 + 2x_2 - 5x_3 \end{bmatrix} = f(\boldsymbol{x})$$

Exercises

1. Find the matrix representation of each of the following linear functions:

a) $f(x_1, x_2) = (2x_1 - x_2,\ x_1 + 4x_2,\ -7x_1,\ 3x_1 + 5x_2)$.

b) $\boldsymbol{R}(t) = 4t\boldsymbol{i} - 5t\boldsymbol{j} - 2t\boldsymbol{k}$.

c) $L(x) = 6x$.

2. Let g be define by $g(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{A} = \begin{bmatrix} 2 & -1 \\ -2 & 1 \\ 0 & -3 \\ 3 & 5 \end{bmatrix}$. Find $g(3,-9)$.

3. Let $f : \boldsymbol{R}^2 \to \boldsymbol{R}^2$ be the function in which $f(\mathbf{x})$ is the vector that results from rotating the vector $\boldsymbol{x}$ about the origin $\frac{\pi}{4}$ in the counterclockwise direction.

a)Explain why *f* is a linear function.

b)Find the matrix representation for *f*.

d)Find $f(4,-9)$.

4. Let $f : \boldsymbol{R}^2 \to \boldsymbol{R}^2$ be the function in which $f(\mathbf{x})$ is the vector that results from rotating the vector $\boldsymbol{x}$ about the origin θ in the counterclockwise direction. Find the matrix representation for *f*.

5. Suppose $g : \boldsymbol{R}^2 \to \boldsymbol{R}^2$ is a linear function such that $g(1,2) = (4,7)$ and $g(-2,1) = (2,2)$.

Find the matrix representation of *g*.

6. Suppose $f : \boldsymbol{R}^n \to \boldsymbol{R}^p$ and $g : \boldsymbol{R}^p \to \boldsymbol{R}^q$ are linear functions. Prove that the composition $g \circ f : \boldsymbol{R}^n \to \boldsymbol{R}^q$ is a linear function.

7. Suppose $f : \boldsymbol{R}^n \to \boldsymbol{R}^p$ and $g : \boldsymbol{R}^n \to \boldsymbol{R}^p$ are linear functions. Prove that the function $f + g : \boldsymbol{R}^n \to \boldsymbol{R}^p$ defined by $(f + g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x})$ is a linear function.

6.2 Matrix Algebra

Let us consider the composition $h = g \circ f$ of two linear functions $f : \boldsymbol{R}^n \to \boldsymbol{R}^p$ and $g : \boldsymbol{R}^p \to \boldsymbol{R}^q$. Suppose $\boldsymbol{A}$ is the matrix of *f* and $\boldsymbol{B}$ is the matrix of *g*. Let's see about the matrix $\boldsymbol{C}$ of *h*. We know the columns of $\boldsymbol{C}$ are the vectors $g(f(\boldsymbol{e}_j)), j = 1,2,\ldots,n$, where, of course, the vectors $\boldsymbol{e}_j$ are the coordinate vectors for $\boldsymbol{R}^n$. Now the columns of $\boldsymbol{A}$ are just the vectors $f(\boldsymbol{e}_j), j = 1,2,\ldots,n$. Thus the vectors $g(f(\boldsymbol{e}_j))$ are simply the products $\boldsymbol{B}f(\boldsymbol{e}_j)$. In other words, if we denote the columns of $\boldsymbol{A}$ by $\boldsymbol{k}_i$, $i = 1,2,\ldots,n$, so that $\boldsymbol{A} = [\boldsymbol{k}_1, \boldsymbol{k}_2, \ldots, \boldsymbol{k}_n]$, then the columns of $\boldsymbol{C}$ are $\boldsymbol{Bk}_1, \boldsymbol{Bk}_2, \ldots, \boldsymbol{Bk}_n$, or in other words, $\boldsymbol{C} = [\boldsymbol{Bk}_1, \boldsymbol{Bk}_2, \ldots, \boldsymbol{Bk}_n]$.

Example

Let the matrix $\boldsymbol{B}$ of g be given by $\boldsymbol{B} = \begin{bmatrix} 1 & 0 & 2 \\ -1 & -5 & 8 \\ 2 & 7 & -3 \\ 2 & -2 & 1 \end{bmatrix}$ and let the matrix $\boldsymbol{A}$ of f be given by $\boldsymbol{A} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \\ -4 & -3 \end{bmatrix}$. Thus $f: \boldsymbol{R^2} \to \boldsymbol{R^3}$ and $g: \boldsymbol{R^3} \to \boldsymbol{R^4}$ (Note that for the composition $h = g \circ f$ to be defined, it must be true that the number of columns of $\boldsymbol{B}$ be the same as the number of rows of $\boldsymbol{A}$.). Now, $\boldsymbol{k}_1 = \begin{bmatrix} 3 \\ 1 \\ -4 \end{bmatrix}$ and $\boldsymbol{k}_2 = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$, and so

$\mathbf{Bk}_1 = \begin{bmatrix} -5 \\ -40 \\ 25 \\ 0 \end{bmatrix}$ and $\mathbf{Bk}_2 = \begin{bmatrix} -5 \\ -35 \\ 25 \\ -3 \end{bmatrix}$. The matrix $\boldsymbol{C}$ of the composition is thus

$$\mathbf{C} = \begin{bmatrix} -5 & -5 \\ -40 & -35 \\ 25 & 25 \\ 0 & -3 \end{bmatrix}.$$

These results inspire us to define a product of matrices. Thus, if $\boldsymbol{B}$ is an $n \times p$ matrix, and $\boldsymbol{A}$ is a $p \times q$ matrix, the ***product*** $\boldsymbol{BA}$ of these matrices is defined to be the $n \times q$ matrix whose columns are the column vectors $\boldsymbol{Bk}_j$, where $\boldsymbol{k}_j$ is the j^{th} column of $\boldsymbol{A}$. Now we can simply say that the matrix representation of the composition of two linear functions is the product of the matrices representing the two functions.

There are several interesting and important things to note regarding matrix products. First and foremost is the fact that in general $\boldsymbol{BA} \neq \boldsymbol{AB}$, even when both products are defined (The product ***BA*** obviously defined only when the number of columns of ***B*** is the same as the number of rows of ***A***.). Next, note that it follows directly from the fact that $h \circ (f \circ g) = (h \circ f) \circ g$ that for ***C***(***BA)*** = (***CB***)***A***. Since it does not matter where we insert the parentheses in a product of three or more matrices, we usually omit them entirely.

It should be clear that if *f* and *g* are both functions from $\boldsymbol{R}^n$ to $\boldsymbol{R}^p$, then the matrix representation for the sum $f + g\colon \boldsymbol{R}^n \to \boldsymbol{R}^p$ is simply the matrix

$$\boldsymbol{A} + \boldsymbol{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & & & \\ a_{p1} + b_{p1} & a_{p2} + b_{p2} & \dots & a_{pn} + b_{pn} \end{bmatrix},$$

where

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}$$

is the matrix of *f*, and

$$\boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{bmatrix}$$

is the matrix of g. Meditating on the properties of linear functions should convince you that for any three matrices (of the appropriate sizes) $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$, it is true that

$$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB}+\boldsymbol{AC}.$$

Similarly, for appropriately sized matrices, we have $(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C} = \boldsymbol{AC}+\boldsymbol{BC}$.

Exercises

8. Find the products:

a) $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ b) $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix}$

c) $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ 1 & 3 \end{bmatrix}$ d) $\begin{bmatrix} 1 & -3 & 2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 5 \\ -2 & 3 \\ 0 & 2 \\ -3 & 4 \end{bmatrix}$

9. Find a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ b) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

10. Let $\boldsymbol{A}(\theta)$ be the 2×2 matrix for the linear function that rotates the plane θ counterclockwise. Compute the product $\boldsymbol{A}(\theta)\boldsymbol{A}(\eta)$, and use the result to give identities for $\cos(\theta+\eta)$ and $\sin(\theta+\eta)$ in terms of $\cos\theta$, $\cos\eta$, $\sin\theta$, and $\sin\eta$.

11. a)Find the matrix for the linear function that rotates $\boldsymbol{R}^3$ about the coordinate vector $\boldsymbol{j}$ by $\frac{\pi}{4}$ (In the positive direction, according to the usual “right hand rule” for rotation.).

b)Find a vector description for the curve that results from applying the linear transformation in a) to the curve $\boldsymbol{R}(t) = \cos t\boldsymbol{i} + \sin t\boldsymbol{j} + t\boldsymbol{k}$.

12. Suppose $f:\boldsymbol{R}^2 \to \boldsymbol{R}^2$ is linear. Let C be the circle of radius 1 and center at the origin. Find a vector description for the curve $f(C)$.

13. Suppose $g:\boldsymbol{R}^2 \to \boldsymbol{R}^n$ is linear. Suppose moreover that $g(1,1) = (2,3)$ and $g(-1,1) = (4,-5)$. Find the matrix of g.