

# Interlude: Abstract Linear Transformations and Vector Spaces

In this interlude, we take a step back from matrix algebra and give a more general setting to the linear problems addressed in this book. By “linear problems” we mean problems involving equations consisting of linear combinations of variables. These problems have come in two general forms: (i) solving a system of linear equations  $\mathbf{Ax} = \mathbf{b}$ , and (ii) describing the behavior of iterative models of the form  $\mathbf{p}' = \mathbf{Ap}$ .

In Section 4.1 we used the term *linear transformation* to refer to mappings  $T : \mathbf{w} \rightarrow \mathbf{w}' = T(\mathbf{w})$ , where  $\mathbf{w}' = \mathbf{Aw}$ . In that section we treated the system  $\mathbf{w}' = \mathbf{Aw}$  as defining a computer graphics transformation  $T(\mathbf{w})$  that might be applied to a stick-figure drawing or possibly the whole  $x$ - $y$  plane (or  $x$ - $y$ - $z$  space). The key property of linear transformations is (Theorem 1 of Section 4.1)

$$T(a\mathbf{w} + b\mathbf{v}) = aT(\mathbf{w}) + bT(\mathbf{v}) \quad (1)$$

Property (1) led to the observation that linear transformations take lines into lines.

Property (1) turns out to be at the heart of all linear models. If a transformation  $T(\mathbf{w})$  satisfies (1), where  $\mathbf{w}$  is an  $n$ -vector, then  $T(\mathbf{w})$  must actually be a matrix transformation:  $T(\mathbf{w}) = \mathbf{Aw}$ . That is, for vectors, property (1) defines a matrix-vector product.

We define an **abstract linear transformation** on  $n$ -vectors to be any mapping  $\mathbf{w}' = T(\mathbf{w})$  of an  $n$ -vector  $\mathbf{w}$  to an  $m$ -vector  $\mathbf{w}'$  such that  $T$  satisfies (1).

**Theorem 1.** Any abstract linear transformation  $\mathbf{w}' = T(\mathbf{w})$  can be represented by matrix multiplication:  $\mathbf{w}' = \mathbf{Aw}$ .

*Proof:* If  $\mathbf{w}$  is an  $n$ -vector and  $\mathbf{w}'$  is an  $m$ -vector, then  $\mathbf{A}$  will have to be an  $m$ -by- $n$  matrix. When transforming a set of points in Section 4.1, we used property (1) to simplify the calculations by expressing

all points as linear combinations of a few key points. In this proof we reason the same way. Here the key points will be unit vectors, such as  $(1, 0, 0, \dots, 0)$ . Let  $\mathbf{e}_j$  be the  $j$ th unit vector in  $n$  dimensions (with a 1 in the  $j$ th position and 0's in the other  $n - 1$  positions).

We define the  $j$ th column  $\mathbf{a}_j^C$  of  $\mathbf{A}$  to be  $T(\mathbf{e}_j)$ , the vector to which  $\mathbf{e}_j$  is mapped by  $T$ :

$$\mathbf{a}_j^C = T(\mathbf{e}_j) \quad (2)$$

If  $\mathbf{w} = (w_1, w_2, \dots, w_n)$ , we can write  $\mathbf{w}$  as

$$\mathbf{w} = w_1\mathbf{e}_1 + w_2\mathbf{e}_2 + \dots + w_n\mathbf{e}_n \quad (3)$$

Then by (1) and (2)

$$T(\mathbf{w}) = w_1T(\mathbf{e}_1) + w_2T(\mathbf{e}_2) + \dots + w_nT(\mathbf{e}_n) \quad (4a)$$

$$= w_1\mathbf{a}_1^C + w_2\mathbf{a}_2^C + \dots + w_n\mathbf{a}_n^C \quad (4b)$$

$$= \mathbf{A}\mathbf{w}$$

The linear combination in (4b) of the columns of  $\mathbf{A}$  is exactly the definition of  $\mathbf{A}\mathbf{w}$ . ■

The preceding proof shows that any linear transformation is specified by knowing what it does with a set of coordinate vectors. In the proof we used the unit vectors  $\mathbf{e}_j$ , but any set of vectors whose linear combinations yield all other vectors would work.

In Section 2.5 we saw that for square matrices, eigenvector coordinates made matrix multiplication very easy. If  $\mathbf{u}_i$  is an eigenvector of  $T$ , then  $T(\mathbf{u}_i) = \lambda_i\mathbf{u}_i$ . In the proof above, if  $\mathbf{w}$  has the representation in eigenvector coordinates of

$$\mathbf{w} = r_1\mathbf{u}_1 + r_2\mathbf{u}_2 + \dots + r_n\mathbf{u}_n \quad (5)$$

then (4a) and (4b) become

$$\begin{aligned} T(\mathbf{w}) &= r_1T(\mathbf{u}_1) + r_2T(\mathbf{u}_2) + \dots + r_nT(\mathbf{u}_n) \\ &= \lambda_1r_1\mathbf{u}_1 + \lambda_2r_2\mathbf{u}_2 + \dots + \lambda_nr_n\mathbf{u}_n \end{aligned} \quad (6)$$

For example, the rabbit-fox growth model (Example 5 of Section 3.1)

$$\mathbf{p}' = \mathbf{A}\mathbf{p}, \quad \text{where } \mathbf{A} = \begin{bmatrix} 1.1 & -.15 \\ .1 & .85 \end{bmatrix} \quad (7)$$

has eigenvectors  $\mathbf{u}_1 = [3, 2]$  and  $\mathbf{u}_2 = [1, 1]$  with associated eigenvalues  $\lambda_1 = 1$  and  $\lambda_2 = .95$ . So if the initial population  $\mathbf{p} = [R, F]$  were written in eigenvector coordinates as  $\mathbf{p} = [s_1, s_2]$  ( $= s_1\mathbf{u}_1 + s_2\mathbf{u}_2$ ), then the linear transformation  $T$  given by  $\mathbf{A}$  becomes

$$\begin{aligned}
 T(\mathbf{p}) &= s_1T(\mathbf{u}_1) + s_2T(\mathbf{u}_2) \quad \text{or} \\
 \mathbf{T} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 0 & .95 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} s_1 \\ .95s_2 \end{bmatrix} \\
 &= s_1(1\mathbf{u}_1) + s_2(.95\mathbf{u}_2)
 \end{aligned} \tag{8}$$

The concept of a linear transformation helps to remind us that (7) and (8) are the same thing—that is, are the same linear transformation—but expressed in different coordinate systems.

The lesson from Section 2.5 is that, when possible, linear transformations should be expressed in eigenvector coordinates. To convert to eigenvector coordinates and afterwards to convert back to standard coordinates, we can use the matrix equation (Theorem 5 of Section 3.3)

$$\mathbf{A} = \mathbf{U}\mathbf{D}_\lambda\mathbf{U}^{-1} \tag{9}$$

The concept of a linear transformation also gives new understanding to the problem of solving a system of equations. Consider our refinery problem:

$$\begin{aligned}
 \mathbf{Ax} = \mathbf{b}: \quad \text{Heating oil:} \quad & 20x_1 + 4x_2 + 4x_3 = 500 \\
 \text{Diesel oil:} \quad & 10x_1 + 14x_2 + 5x_3 = 850 \\
 \text{Gasoline:} \quad & 5x_1 + 5x_2 + 12x_3 = 1000
 \end{aligned} \tag{10}$$

Viewing (10) as a linear transformation problem,

$$T(\mathbf{x}) = \mathbf{b} \tag{11}$$

we see that solving (11) for a vector  $\mathbf{x}$  of production levels is asking for a vector  $\mathbf{x}$  in the domain of  $T$  that is mapped by  $T$  to  $\mathbf{b}$ . *This is a vector-valued version of the problem: Given a function  $f(x)$  and a constant  $b$ , find an  $x$  for which  $f(x) = b$ .*

Just like the function version of this problem, if  $\mathbf{b}$  is in the range of  $T$ , there will be at least one solution; if  $T$  is a one-to-one mapping, there will be at most one solution; otherwise, there may be many solutions.

Linear transformations provide a convenient way to abstract matrix problems. However, matrix problems are only the beginning. Linear transformations can be defined to act on more complicated sets than vectors, such as functions.

Functions can be thought of as an infinite dimensional extension of vectors. An **abstract vector space** is any collection  $\mathbf{C}$  of elements that obey the law of linearity. That is, if  $A$  and  $B$  are elements of  $\mathbf{C}$ , then  $rA + sB$  are in  $\mathbf{C}$ , for any constants  $r, s$ . The set of all continuous functions forms an abstract vector space. The same is true for the set of all functions, continuous or not.

### Example 1. Linear Transformations on Functions

- (i) *Shift transformation*  $S_a$ . Let  $S_a(f(x)) = f(x + a)$ . That is,  $S_5$  shifts the values of a function  $f(x)$  5 units to the left:

$$\text{if } f(x) = x^2 - 2x, \quad \text{then } S_5(f(x)) = (x + 5)^2 - 2(x + 5)$$

- (ii) *Reflection transformation*  $R$ . Let  $R(f(x)) = f(-x)$ . So  $R$  reflects the graph of any function about the  $y$ -axis.
- (iii) *Differentiation transformation*  $D$ . Let  $D(f(x)) = df(x)/dx$ , the derivative of  $f(x)$  (assuming that the derivative exists).
- (iv) *Integration transformation*  $I$ . Let  $I(f(x)) = \int f(x) dx$ , the integral of  $f(x)$  (integration actually requires a constant term; here we will assume that the constant is 0). ■

It is left to the reader to check that the transformations in Example 1, parts (i) and (ii) are indeed linear transformations. The required property, generalizing (1), is

$$T(af(x) + bg(x)) = aT(f(x)) + bT(g(x)) \quad \text{for any constants } a, b \quad (12)$$

Because differentiation is so important, let us check that it is a linear transformation. Property (12) is

$$\frac{d}{dx} [af(x) + bg(x)] = a \frac{d}{dx} f(x) + b \frac{d}{dx} g(x) \quad (13)$$

But (13) is the linearity rule of derivatives. Similarly, for integration we have

$$\int [af(x) + bg(x)] dx = a \int f(x) dx + b \int g(x) dx \quad (14)$$

Virtually all of the theory for analyzing matrix equations extends to linear transformations of functions. For example, we can talk about inverse transformations and about eigenfunctions  $u(x)$ :  $T(u(x)) = \lambda u(x)$ .

### Example 2. Inverse Transformations of Linear Transformations

- (i) For the shift transformation  $S_a(f(x)) = f(x + a)$ , the inverse transformation  $S^{-1}$  is  $S_{-a}$  since  $S_{-a}[S_a(f(x))] = f(x + a - a) = f(x)$ .
- (ii) For the reflection transformation  $R(f(x)) = f(-x)$ , the inverse  $R^{-1}$  is simply the reflection transformation itself. So  $R^{-1} = R$ .
- (iii) For differentiation, there is no (unique) inverse  $D^{-1}$ . If two functions differ by a constant, say,  $x^2 + x + 2$  and  $x^2 + x + 5$ , they have the same derivative,  $2x + 1$ . So  $D^{-1}(2x + 1)$  cannot be uniquely defined.
- (iv) For integration, the inverse  $I^{-1}$  is differentiation. That is,

$$\frac{d}{dx} [\int f(x) dx] = f(x)$$

This is the fundamental theorem of calculus! ■

### Example 3. Eigenfunctions of Linear Transformations

- (i) For the shift transformation  $S_a(f(x)) = f(x + a)$ , an eigenfunction  $u(x)$  must have the property that  $u(x + a) = u(x)$ . For example, when  $a = 2\pi$ , the trigonometric functions, such as  $\sin x$  or  $\cos x$ , are eigenfunctions of  $S_{2\pi}$ , with eigenvalue 1.
- (ii) For the reflection transformation  $R(f(x)) = f(-x)$ , an eigenfunction  $u(x)$  associated with  $\lambda = 1$  is any symmetric  $u(x)$ , that is,  $u(x) = u(-x)$ .
- (iii) For differentiation,  $e^{\lambda x}$  is the eigenfunction of  $D$  associated with eigenvalue  $\lambda = k$ , since  $de^{kx}/dx = ke^{kx}$ .
- (iv) For integration  $e^x/\lambda$  is the eigenfunction of  $I$  associated with eigenvalue  $\lambda = k$ , since integration is the reverse operation of differentiation. ■

There is one very important generalization of differentiation that bears special mention, namely differential equations. The differential equation

$$y''(x) - 2y'(x) = f(x) \tag{15}$$

can be considered a linear transformation DE of  $y(x)$  to  $f(x)$ , that is,  $DE(y(x)) = f(x)$ . It is left to the reader to check that DE satisfies property (1). Any differential equation whose left side is a linear combination of derivatives will be a linear transformation. The advanced theory of differential equations is based heavily on eigenfunctions and inverse transformations.

For linear transformations defined in terms of matrices, we noted that the “right” coordinates for describing the transformation are eigenvector coordinates. The same applies to linear transformations of functions. Functions should be expressed as linear combinations (infinite series) of eigenfunctions of the linear transformation.

This book is not about linear transformations of functions. It is about matrices and vectors. But it is important to be aware of the powerful generalizations of matrices and matrix algebra which are the basis for much of higher mathematics. If the reader masters the matrix-based linear algebra in this book, he or she will have an excellent foundation for any future work with functional linear algebra.

The purpose of this interlude has been to implant the seed in the reader’s mind that many operations on functions are linear transformations and that most of the theory of matrices extends to these linear transformations. In Chapter 5, occasional examples using linear transformations of functions are given.

