Lesson 14

Reading: Larson, Section 5.1, Length and Dot Product in R^n (partial; see below); Larson, Section 7.3, Symmetric Matrices and Orthogonal Diagonalization

Suggested exercises: Larson, Section 5.1: 3, 9, 11, 45; Larson, Section 7.3: 1, 3, 7, 11, 13, 23, 27, 33, 35, 41, 43, 49

Submit: Lesson 14: Symmetric matrices

Section 5.1: Length and Dot Product in R^n

This section is not meant to be covered thoroughly. We are interested only in the basic idea of the dot product, and its use in computing lengths of vectors and the angles between them. The goal is to familiarize yourself with dot products enough to use them in working with orthogonal matrices.

In your reading, you can ignore the Cauchy-Schwarz inequality, the triangle inequality, and the Pythagorean theorem. Chances are good that you already have experience with the dot product: if so, this is a chance to review. There are only a few suggested exercises from this section.

The dot product of a pair of vectors $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ in R^n is

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

The dot product of two vectors can also be shown to be equal to the product of the lengths of the two vectors and the cosine of the angle between them:

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

where $\|\mathbf{v}\|$ denotes the length of a vector $\mathbf{v}$, and we always take $0° \leq \theta \leq 180°$.

This gives us the property that nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ are *orthogonal* (another word for perpendicular) if $\mathbf{x} \cdot \mathbf{y} = 0$, because nonzero vectors have nonzero length and therefore the cosine of the angle between them must be zero, which means the angle is $90°$.

Note that this gives us a way to compute the length of a vector $\mathbf{v}$. Since the angle between a nonzero vector and itself is $0°$ and $\cos 0° = 1$, we have $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$, and therefore, the length of $\mathbf{v}$ is

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

If $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, then

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

For example, in R^3, the length of the vector $\mathbf{v} = (1, -1, 0)$ is

$$\|\mathbf{v}\| = \sqrt{1^2 + (-1)^2 + 0^2} = \sqrt{2}.$$

If we divide a vector $\mathbf{v}$ by its length, giving us

$$\frac{1}{\|\mathbf{v}\|}\mathbf{v},$$

then we get a *unit vector* (a vector with length 1) in the direction of $\mathbf{v}$. We say that we have *normalized* $\mathbf{v}$. For example, we can normalize the vector $(1, -1, 2)$ in R^3, giving us the unit vector

$$\frac{1}{\sqrt{1^2 + (-1)^2 + 2^2}}(1, -1, 2) = \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right).$$

Section 7.3: Symmetric Matrices and Orthogonal Diagonalization

In section 7.2, we saw that some matrices can be diagonalized, but not all. In this section, we will focus on an important class of matrices whose members can all be diagonalized: the symmetric matrices.

A *symmetric* matrix, you will recall, is one that is equal to its transpose. The symmetry referred to in the name is a symmetry in the components of the matrix with respect to the main diagonal: components symmetrically opposite the main diagonal must be the same. In other words, for an $n \times n$ symmetric matrix A whose component in row i, column j is denoted by a_{ij}, we must have $a_{ij} = a_{ji}$ for all values of i, j from 1 to n.

Thus, for example,

$$A = \begin{bmatrix} 1 & 0 & -2 & 3 \\ 0 & -2 & -1 & 6 \\ -2 & -1 & 3 & 5 \\ 3 & 6 & 5 & 0 \end{bmatrix}$$

is symmetric. Each element a_{ij} is the same as the a_{ji} element for $1 \leq i, j \leq 4$. Each row is equal to the corresponding column of the matrix.

Symmetric matrices have a number of important properties. First, all of their eigenvalues must be real. In this chapter, the author has focused on matrices with real eigenvalues, but you should be aware that a matrix with real entries can have complex eigenvalues (assuming we are considering our vector space as having the set of complex numbers as its scalars). For example, the matrix

$$A = \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$$

has characteristic equation

$$(\lambda - 2)^2 + 3^3 = 0,$$

and therefore has eigenvalues $\lambda = 2 \pm 3i$. The eigenvalue of $\lambda = 2 + 3i$ has eigenvector

$$\begin{bmatrix} i \\ 1 \end{bmatrix},$$

and the eigenvalue $\lambda = 2 - 3i$ has eigenvector

$$\begin{bmatrix} -i \\ 1 \end{bmatrix}.$$

(The eigenvectors are found by the usual means; the only thing that is new is that complex numbers show up as coefficients in the linear systems and therefore in the solutions.)

If a matrix is symmetric, however, this cannot occur: the eigenvalues must be real. Furthermore, for an $n \times n$ matrix, we are guaranteed to have n linearly independent eigenvectors. This does not hold in general; for example, the matrix

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

has a repeated eigenvalue $\lambda = 2$, with multiplicity 2, and only one linearly independent eigenvector,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

This matrix cannot be diagonalized. (Note that it is not symmetric.)

Unlike the example just given, for a symmetric matrix, if the multiplicity of an eigenvalue λ is m, then there must be m associated linearly independent eigenvectors of λ.

For example, the symmetric matrix

$$A = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

has characteristic equation

$$(\lambda - 4)^2(\lambda - 2)^2 = 0,$$

and therefore has eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 2$, each with multiplicity 2. The eigenvalue $\lambda_1 = 4$ has linearly independent eigenvectors $(0, 0, 1, 0)$ and $(1, 1, 0, 0)$; the eigenvalue $\lambda_2 = 2$ has linearly independent eigenvectors $(0, 0, 0, 1)$ and $(-1, 1, 0, 0)$.

Note that the entire set of eigenvectors is linearly independent, so A has a set of 4 linearly independent eigenvectors, which means it can be diagonalized. This is true of all symmetric matrices: an $n \times n$ symmetric matrix has n linearly independent eigenvectors, and therefore can be diagonalized.

There is a specific type of matrix that can diagonalize a symmetric matrix, called an *orthogonal* matrix. An orthogonal matrix is a matrix whose transpose is equal to its inverse. That is, P is orthogonal if $P^{-1} = P^T$. A simple example of an orthogonal matrix is

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

There is another way of characterizing orthogonal matrices in terms of the columns of the matrix.

A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is said to be an *orthogonal set* of vectors if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ whenever $i \neq j$. For example, the set of eigenvectors in the preceding example,

$$\{(0, 0, 1, 0), (1, 1, 0, 0), (0, 0, 0, 1), (-1, 1, 0, 0)\},$$

forms an orthogonal set. You should check that the dot products between different vectors in this set are all zero.

If we insist not only that the dot products between different vectors is zero, but also that every vector in the set has length 1, then the set is called an *orthonormal* set of vectors. In other words, a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is orthonormal if

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \text{ if } i \neq j$$

for $1 \leq i, j \leq n$, and

$$\mathbf{v}_i \cdot \mathbf{v}_i = 1$$

for $1 \leq i \leq n$.

Any orthogonal set of vectors can be turned into an orthonormal set by normalizing the vectors in the set, that is, by dividing each vector by its length. Thus, the matrix in the preceding example has an orthonormal set of eigenvectors:

$$\left\{(0,0,1,0), \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right), (0,0,0,1), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)\right\}.$$

This set of eigenvectors can be used to diagonalize the matrix A: if we let

$$P = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

then

$$P^{-1} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \end{bmatrix},$$

which you can see is also P^T, so our matrix is an orthogonal matrix. This is no coincidence: whenever the columns of a matrix form an orthonormal set, the matrix is orthogonal.

Now, finally, we can state our other characterization of orthogonal matrix: P is an orthogonal matrix if and only if its columns form an orthonormal set.

We now have enough machinery to state, definitively, the diagonalization property of symmetric matrices: if an $n \times n$ matrix A is symmetric, then A has n real eigenvalues (counting multiplicity), with n linearly independent eigenvectors that form an orthonormal set, which gives us an orthogonal matrix P with this orthonormal set of eigenvectors as its columns (and for which $P^{-1} = P^T$), and this matrix can be used to diagonalize A:

$$D = P^{-1}AP = P^T AP,$$

where D is a diagonal matrix with the eigenvalues of A on its main diagonal. The matrix A is said to be *orthogonally diagonalizable*.

Symmetric matrices arise frequently in applications; this is why so much time and attention have been devoted to them here.

Let's see an example. Suppose

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

We observe that this matrix is symmetric. The characteristic equation is

$$\begin{vmatrix} \lambda - 1 & 2 \\ 2 & \lambda - 4 \end{vmatrix} = 0,$$

or

$$\lambda^2 - 5\lambda = 0,$$

which produces eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 5$. The eigenvalue $\lambda_1 = 0$ has eigenvector satisfying the homogeneous system with coefficient matrix

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

giving an eigenvector $\mathbf{v}_1 = (2, 1)$. The eigenvalue $\lambda_2 = 5$ has an eigenvector satisfying the homogeneous system with coefficient matrix

$$\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix},$$

giving eigenvector $\mathbf{v}_2 = (1, -2)$. Note that the dot product of these eigenvectors is 0, so they are orthogonal. Divide each of the eigenvectors by its length, giving the orthonormal pair

$$\left\{ \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right), \left(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}} \right) \right\}.$$

The matrix P is therefore

$$P = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix},$$

and since its columns come from an orthonormal set, we know that

$$P^{-1} = P^T = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix}.$$

(You may have noticed that the inverse is equal to P in this case. This won't always happen, of course.) Therefore, we have the orthogonal diagonalization

$$D = P^T AP = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix}.$$

A final observation. This relatively pleasant property of being orthogonally diagonalizable is one that is enjoyed by the symmetric matrices, and *only* by the symmetric matrices. There are, of course, non-symmetric matrices that can be diagonalized, but they will not be diagonalized by orthogonal matrices.