Homework 6

Monday, October 19, 2020 10:13 PM

document

Chapter 17

Homework 6: Eigenvalues and Eigenvectors

Contents

Learning Objectives

Concepts

- Compute the eigenvalues and eigenvectors of a 2×2 matrix by hand
- Compute the eigenvalues and eigenvectors of an $n \times n$ matrix using MATLAB
- Describe the geometric meaning of eigenvalues and eigenvectors
- Use eigenvectors to compute and interpret directions of variation in data

MATLAB skills

- Compute the eigenvectors and eigenvalues of a given matrix
- From a given dataset, set up the relevant matrices and compute the covariance matrix of the dataset.

What is this about? The big ideas of this assignment are eigenvectors and eigenvalues. Recall that when you multiply a vector by a matrix, the resulting vector usually points in a different direction. An eigenvector of a square matrix is a vector which does not change direction when multiplied by that matrix. It can only change in length. The eigenvalue corresponding to this eigenvector is the scale factor that is applied to that eigenvector as a result of the matrix multiplication. Therefore, the eigenvector of a matrix points in a special direction – its a direction that is not modified by the linear transformation associated with that matrix. This is an idea that we will keep coming back to in a number of different ways throughout QEA (including next semester). The ideas contained here can be applied in many ways (many of which we won't get to until next semester) such as

- Directions of greatest variation in data.

- Natural co-ordinates of systems.
- Frequency response of filters.
- Analysis of dynamical systems.

Reference Material Here are some videos and tutorials that may help you understand this material.

- Eigenvalues and Eigenvectors by 3Blue1Brown (watch first 14 mins)
- Paul's Online Notes. Review : Eigenvalues and Eigenvectors
- Intro to eigenvectors by PatrickJMT
- Calculating eigenvalues and eigenvectors of a 2×2 matrix by PatrickJMT.

17.1 Calculating Eigenvalues and Eigenvectors of Matrices

Recall from class that λ is an eigenvalue of a matrix $\mathbf{A}$ with corresponding eigenvector $\mathbf{v}$ if $\mathbf{Av} = \lambda\mathbf{v}$. Geometrically, this means that the matrix $\mathbf{A}$ doesn't change the direction of $\mathbf{v}$, it simply scales it by a factor of λ.

Given a square matrix, how can we find its eigenvalues and eigenvectors? In class, we calculated these by hand for the special case of diagonal matrices, and now we will move to generic 2×2 matrices. For general square matrices which are larger than 2×2, we will use MATLAB's `eig` to compute the eigenvalues and eigenvectors.

Finding eigenvalues

So far we've dealt with matrices for which it is possible to think your way to the eigenvalues. For general matrices, this is rarely the case, and we need a method that is foolproof. The method most widely adopted involves the determination of an algebraic equation for the eigenvalues, usually known as the *characteristic* equation. For this reason, eigenvalues are often known as *characteristic* values.

Let's start with an example. Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 18 & -2 \\ 12 & 7 \end{bmatrix}$$

The definition of an eigenvalue and eigenvector imply that we are seeking λ and $\mathbf{v}$ which satisfy

$$\mathbf{Av} = \lambda\mathbf{v}.$$

We subtract $\lambda\mathbf{v}$ from both sides

$$\mathbf{Av} - \lambda\mathbf{v} = \mathbf{0}$$

and then factor the left hand side to give

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$$

Notice that an identity matrix $\mathbf{I}$ has appeared out of nowhere - this simply allows us to write the vector $\mathbf{v}$ as $\mathbf{Iv}$ so that we can factor out the matrix $\mathbf{A} - \lambda\mathbf{I}$. Also notice that this new matrix is just $\mathbf{A}$ with λ subtracted from the diagonal terms. For this example we have

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 18-\lambda & -2 \\ 12 & 7-\lambda \end{bmatrix}.$$

We are only interested in $\mathbf{v}$ that are nonzero, i.e., $\mathbf{v}$ is not the vector of all zeroes. (This is because $\mathbf{v} = \mathbf{0}$ is always a solution to $\mathbf{Av} = \lambda\mathbf{v}$ for any $\mathbf{A}$ and any λ, so it's not very interesting or informative.) Assuming $\mathbf{v}$ is nonzero implies that the matrix $(\mathbf{A} - \lambda\mathbf{I})$ is not invertible. Why? If $(A - \lambda\mathbf{I})$ were invertible, then we could rearrange the equation to get

$$\underbrace{\mathbf{v} = (\mathbf{A} - \lambda\mathbf{I})^{-1}\mathbf{0}} = \mathbf{0}$$

Divide by $(A - \lambda I)$

which contradicts our assumption that $\mathbf{v} \neq \mathbf{0}$. THerefore, $(\mathbf{A} - \lambda\mathbf{I})$ is not invertible.

Since $(\mathbf{A} - \lambda\mathbf{I})$ is not invertible, it must have determinant zero. In other words,

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

In our example, this implies that

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (18 - \lambda)(7 - \lambda) + 24 = 0.$$

This is called the characteristic equation:

$$(18 - \lambda)(7 - \lambda) + 24 = 0$$

or, rearranged,

$$\lambda^2 - 25\lambda + 150 = 0.$$

The characteristic equation is a polynomial with the variable λ that arises by setting the determinant of $(\mathbf{A} - \lambda\mathbf{I})$ equal to zero. The solutions to this polynomial give the eigenvalues λ. In our example, the polynomial can be factored

$$(\lambda - 15)(\lambda - 10) = 0$$

so that gives eigenvalues $\lambda_1 = 10$ and $\lambda_2 = 15$. (We could use the quadratic formula if necessary.)

Let's retrace our steps: If λ is either 10 or 15, then the determinant of $(\mathbf{A} - \lambda\mathbf{I})$ is zero. This implies ad-cb that $(\mathbf{A} - \lambda\mathbf{I})$ is not invertible, so we can look for nonzero solutions $\mathbf{v}$ to $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$ and those $\mathbf{v}$ are eigenvectors associated to the eigenvalue λ.

In summary, here's the general procedure for finding the eigenvalues of a matrix:

1. Rearrange $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ to get $(\mathbf{A} - \lambda\mathbf{I}) = \mathbf{0}$.
2. Compute the determinant of $(\mathbf{A} - \lambda\mathbf{I})$.
3. Since the matrix is not invertible, we set that determinant equal to zero: $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. This gives a polynomial in λ, known as the characteristic equation.
4. Solve the polynomial for the roots λ. Those are the eigenvalues.

Exercise 17.1

1. You already know that the eigenvalues of a diagonal matrix are just the entries on the diagonal. Using the above procedure, confirm that

$$\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$$

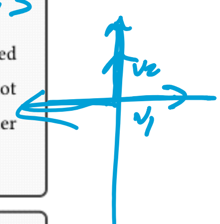
$A - \lambda I = \begin{bmatrix} 2-\lambda & 0 \\ 0 & -3-\lambda \end{bmatrix}$

$\det(A - \lambda I) = (2-\lambda)(3-\lambda) = 0$

$\Rightarrow \lambda = -2, 3$

has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = -3$.

2. Notice that one of the eigenvalues is positive and one is negative. The eigenvector associated with $\lambda_1 = 2$ is $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and the eigenvector associated with $\lambda_2 = -3$ is $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Plot $\mathbf{v}_1, \mathbf{v}_2, \mathbf{A}\mathbf{v}_1$ and $\mathbf{A}\mathbf{v}_2$. What affect does the negative sign in the eigenvalue have? In other words, what is the difference between a negative and positive eigenvalue?



Reverses the direction

It's worth noting that eigenvalues come in more flavors than positive or negative. They can also be complex numbers. For now, will focus on matrices with real eigenvalues, but if you're curious about the complex case, you can learn about it in this worksheet (ignore the first page).

Finding Eigenvectors

In the example in the previous section, we discovered that the eigenvalues of

$$\mathbf{A} = \begin{bmatrix} 18 & -2 \\ 12 & 7 \end{bmatrix}$$

are $\lambda_1 = 10$ and $\lambda_2 = 15$. How do we find the corresponding eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$?

First, let's find the eigenvector corresponding to $\lambda_1 = 10$. Remember that we knew λ_1 was an eigenvalue because it solved the characteristic equation, i.e., $\det(\mathbf{A} - \lambda_1 \mathbf{I}) = \mathbf{0}$. This is important because it implies $(\mathbf{A} - \lambda_1 \mathbf{I})$ is non-invertible, and therefore, there exists a nonzero vector $\mathbf{v}_1$ such that $(\mathbf{A} - \lambda_1 \mathbf{I})\mathbf{v}_1 = \mathbf{0}$. But it's not enough just to know that such a vector exists, we want to know exactly what it is.

In our running example, this means we are looking for $\mathbf{v}_1$ such that

$$(\mathbf{A} - \lambda_1 \mathbf{I})\mathbf{v}_1 = \begin{bmatrix} 18-10 & -2 \\ 12 & 7-10 \end{bmatrix} \mathbf{v}_1 = \begin{bmatrix} 8 & -2 \\ 12 & -3 \end{bmatrix} \mathbf{v}_1 = \mathbf{0}.$$

Let's write $\mathbf{v}_1$ in terms of its unknown components

$$\mathbf{v}_1 = \begin{bmatrix} a \\ b \end{bmatrix},$$

to get the matrix equation

$$\begin{bmatrix} 8 & -2 \\ 12 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This gives us two equations

$$8a - 2b = 0 \text{ and } 12a - 3b = 0.$$

But notice that these equations provide the same information: they both imply that $b = 4a$. This is because $(\mathbf{A} - \lambda_1 \mathbf{I})$ is not invertible, so the rows are linear dependent. The system of linear equations implied by $(\mathbf{A} - \lambda_1 \mathbf{I})\mathbf{v}_1 = \mathbf{0}$ has infinitely many solutions of the form $\begin{bmatrix} a \\ 4a \end{bmatrix}$ for any a. Letting $a = 1$, we get $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$.

If we let $a = 5$, we would have the eigenvector $\begin{bmatrix} 5 \\ 20 \end{bmatrix}$. This hints at an important fact about eigenvectors: *we only care about an eigenvector's direction, not its length.* So we could have chosen $\mathbf{v}_1$ to be any vector pointing the same direction as $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ (such as $\begin{bmatrix} 5 \\ 20 \end{bmatrix}$ or $\begin{bmatrix} -2 \\ -8 \end{bmatrix}$). We often speak about "the" eigenvector corresponding to an eigenvalue, but only the direction of the eigenvector is unique, not the length.

Exercise 17.2

We can always check that λ_1 and $\mathbf{v}_1$ are the corresponding eigenvalue and eigenvector for the matrix $\mathbf{A}$ by plugging them into the equation $\mathbf{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ and verifying that it holds.

1. Use this procedure to check that $\lambda_1 = 10$ and $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ are the corresponding eigenvalue and eigenvector for $\mathbf{A} = \begin{bmatrix} 18 & -2 \\ 12 & 7 \end{bmatrix}$.

$$\begin{bmatrix} 18 & -2 \\ 12 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 40 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \cdot 10$$

Exercise 17.3

Using the basic eigenvalue/eigenvector equation

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

$A(cv) = \lambda(cv)$
divide by c

show that if $\mathbf{v}$ is an eigenvector for λ, then $c\mathbf{v}$ is also an eigenvector for λ, where c is any constant.

Exercise 17.4

In our ongoing example we chose $a = 1$ so that the eigenvector is $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$. But we now know that we could scale this vector to be any length. A very common standard is to normalize eigenvectors so that they are unit length, i.e. their length should be 1.

1. Normalize the eigenvector $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ so that it has unit length.

$|v_1| = \sqrt{1^2+4^2} = \sqrt{17}$

$\boxed{\frac{1}{\sqrt{17}}\begin{bmatrix} 1 \\ 4 \end{bmatrix}}$

Exercise 17.5

Continuing the example above, with

$$\mathbf{A} = \begin{bmatrix} 18 & -2 \\ 12 & 7 \end{bmatrix}$$

find the eigenvector that corresponds to the eigenvalue $\lambda_2 = 15$, and then normalize it so that it has unit length.

$A - \lambda I = \begin{bmatrix} 3 & -2 \\ 12 & -8 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = 0 \Rightarrow \begin{matrix} 3a - 2b = 0 \\ 12a - 8b = 0 \end{matrix}$

$a = \frac{2}{3}b$

$b = \frac{3}{2}a$

$\begin{bmatrix} 1 \\ 3/2 \end{bmatrix}$

$\sqrt{1^2 + (3/2)^2}$

$\sqrt{1 + 9/4} \approx 1.80$

$\boxed{\sqrt{\frac{4}{13}}\begin{bmatrix} 1 \\ 3/2 \end{bmatrix}}$

Exercise 17.6

Determine the eigenvalues and eigenvectors of the following 2x2 matrices. Normalize the eigenvectors.

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

$A - \lambda I = \begin{bmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix}$

$\det(\text{''}) = (4-\lambda)(3-\lambda) - 2$

$= 12 - 4\lambda - 3\lambda + \lambda^2 - 2 \Rightarrow \lambda^2 - 7\lambda + 10 = 0 \Rightarrow \lambda = 2, 5$

$\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} 2a + 2b = 0 \\ b = -a \end{matrix} \Rightarrow \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

$\begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} -a + 2b = 0 \\ a - 2b = 0 \\ b = 1/2\, a \end{matrix} \Rightarrow \begin{bmatrix} 2 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$

2. $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$

$A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 3 & 2-\lambda \end{bmatrix}$

$\det(\text{''}) = (1-\lambda)(2-\lambda) - 6$

$\Rightarrow (1-\lambda)(2-\lambda) - 6 = 0$

$= 2 - 3\lambda + \lambda^2 - 6 = 0 \Rightarrow \lambda^2 - 3\lambda - 4 = 0 \Rightarrow \lambda = 4, -1$

$\begin{bmatrix} -3 & 2 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} -3a + 2b = 0 \\ 3a - 2b = 0 \end{matrix}$ $b = \frac{3}{2}a$

$\begin{bmatrix} 2 \\ 3 \end{bmatrix} \Rightarrow \sqrt{13}\begin{bmatrix} 2 \\ 3 \end{bmatrix}$

Exercise 17.7

We have two vectors,

$$\mathbf{n} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \tag{17.1}$$

but

$$\mathbf{z} = \begin{bmatrix} -1 \\ 1.01 \end{bmatrix} \tag{17.2}$$

In other words, the vectors $\mathbf{n}$ and $\mathbf{z}$ point in a very similar direction, but are not perfectly aligned. Now consider a matrix $\mathbf{S}$ given by

$$\mathbf{S} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{17.3}$$

1. On the same axes, plot the vectors $\mathbf{n}$ and $\mathbf{z}$ using MATLAB.
2. Suppose that $\mathbf{n}$ and $\mathbf{z}$ are transformed by $\mathbf{S}$. On the same axes as in the previous part, plot the vectors $\mathbf{Sn}$ and $\mathbf{Sz}$ using MATLAB.
3. Now, we shall see what happens to these vectors under repeated transformations by $\mathbf{S}$. On the same axes as in the previous part, plot the vectors $\mathbf{SSn}$ and $\mathbf{SSz}$ using MATLAB.
4. On the same axes as in the previous part, plot the vectors $\mathbf{SSSn}$ and $\mathbf{SSSz}$ using MATLAB.
5. On the same axes as in the previous part, plot the vectors $\mathbf{SSSSn}$ and $\mathbf{SSSSz}$ using MATLAB.
6. You should find that $\mathbf{n}$ is unaffected by the transformation by $\mathbf{S}$, but $\mathbf{z}$ on the other hand moves farther and farther away. In other words, under repeated transformations by $\mathbf{S}$, $\mathbf{z}$ grew further and further apart from his four friends. Explain what you see in terms of eigenvalues and eigenvectors.

17.2 Properties of Eigenvalues and Eigenvectors

Consider an $n \times n$ matrix $\mathbf{A}$. The characteristic polynomial will be a polynomial of degree n in λ, i.e., it will have the form

$$c_n\lambda^n + \cdots + c_1\lambda + c_0 = 0$$

where c_i are constants. This polynomial will have n roots, although some of those roots might be the same (e.g., both roots of the polynomial $\lambda^2 + 2\lambda + 1 = 0$ are -1, so we say $\lambda_1 = -1$ and $\lambda_2 = -1$.) Since the eigenvalues are the roots of a polynomial then it is possible that some of them will be *complex*, and their corresponding eigenvectors would be complex too.

The following are key properties of the eigenvalues and eigenvectors (some of these are n-dimensional extensions of what you already saw for 2 dimensions).

- An $n \times n$ matrix has n eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, where it is possible that some eigenvalues are equal or complex.
- If the eigenvalues are distinct (none are equal) then the corresponding eigenvectors are linearly independent.

- If a matrix is symmetric, i.e., $\mathbf{A} = \mathbf{A}^T$, then its eigenvalues are real and its eigenvectors are orthogonal to each other.

17.3 Eigenvalues and Eigenvectors using MATLAB

While most of our work on eigenvalues and eigenvectors has focused on 2D vectors and 2×2 matrices, these ideas extend to higher dimensions as well. The eigenvalues and eigenvectors can be found by solving the characteristic polynomial, which quickly gets out of hand or impossible as the size of $\mathbf{A}$ increases. We can instead use the MATLAB `eig` function.

A few words about `eig` are in order. The following command

```
>> [V,D] = eig(A)
```

will return two matrices. The columns of the matrix V are the eigenvectors. D is a diagonal matrix, with the eigenvalues on the diagonal. The first eigenvector is in the first column of V and has a corresponding eigenvalue in the first diagonal entry of D. Each eigenvector is normalized to have a length or magnitude of 1. The eigenvalues will often "appear" to be sorted according to their size, but this is not necessarily true, and is simply an artifact of the algorithm used to compute them. See the documentation in MATLAB for more details.

Exercise 17.8

Use MATLAB's `eig` function to get the eigenvalues and eigenvectors of the following matrices, and compare to your results from earlier exercises.

1. $\mathbf{A} = \begin{bmatrix} 18 & -2 \\ 12 & 7 \end{bmatrix}$
2. $\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$
3. $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$

17.4 Eigenvalues and Eigenvectors in Data Analysis

In the last class you worked on examples involving correlation matrices. Here we will look at covariance matrices, which are related to correlation matrices, except that the entries are not normalized by the standard deviations of the variables. You can think of covariance matrices as measuring the relationship between random quantities, but without normalization. Thus, information about how small or large these data values are will still be preserved in the covariance matrix.

Suppose that we have two different data variables x and y (e.g. corresponding to temperatures in Boston and Sao Paolo), with x_i and y_i being different values in the data set we can define a a matrix $\mathbf{A}$ as follows:

$$\mathbf{A} = \frac{1}{\sqrt{N-1}} \begin{pmatrix} x_1 - \mu_x & y_1 - \mu_y \\ x_2 - \mu_x & y_2 - \mu_y \\ x_3 - \mu_x & y_3 - \mu_y \\ \vdots & \vdots \\ x_N - \mu_x & y_N - \mu_y \end{pmatrix}$$

where μ_x is the mean of the first column, and N is the number of samples (rows). The covariance matrix of x and y is $\mathbf{R} = \mathbf{A}^T\mathbf{A}$. You can think of the entries of this matrix as storing the un-normalized correlations between the temperatures. Because $\mathbf{R}^T = \mathbf{R}$, this matrix is symmetric, and hence its eigenvalues are real and it has orthogonal eigenvectors.

The eigenvectors and eigenvalues of $\mathbf{R}$ tell us something about how the data are distributed. The eigenvector corresponding to the largest eigenvalue of $\mathbf{R}$, which is also called the *principal eigenvector* of $\mathbf{R}$ points in the direction with the largest variation in the data. The eigenvector corresponding to the second largest eigenvalue points in the direction orthogonal to the principal eigenvector in which there is the second largest amount of variation in the data, and so on (if you have more than 2 dimensional data). The square-root of the eigenvalues tells you about the amount of variation there is in each of those directions. Of course when you only have two different variables in the data set, the matrix $\mathbf{R}$ has only 2 orthogonal eigenvectors.

To illustrate, consider Figure 17.1 which shows the centered (mean subtracted) temperatures of Boston vs Sao Paolo. We have also plotted the two eigenvectors, scaled by the square-root of their corresponding eigenvalues, to illustrate the relative variation of the data along the directions of the two eigenvectors. Notice that the principal eigenvector is in the direction of greatest variation in the data. Figure 17.2 is a similar plot with the temperatures of Boston and Washington DC instead.

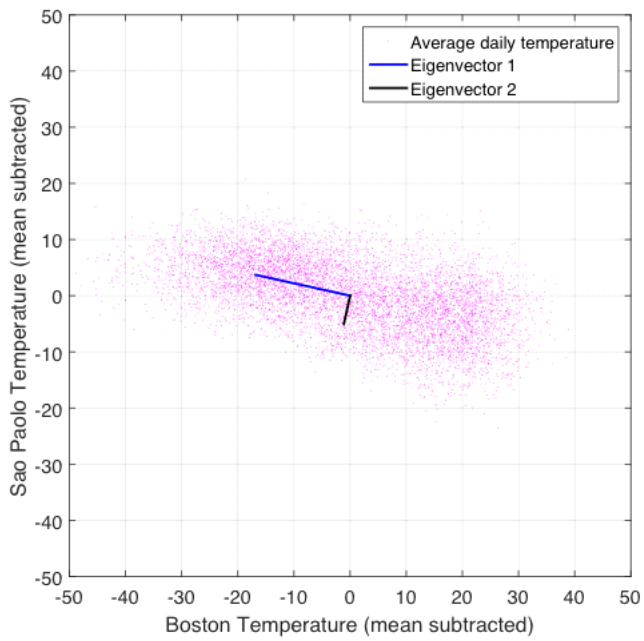


Figure 17.1: Centered average daily temperatures of Boston vs Sao Paolo, with the eigenvectors of the covariance matrix.

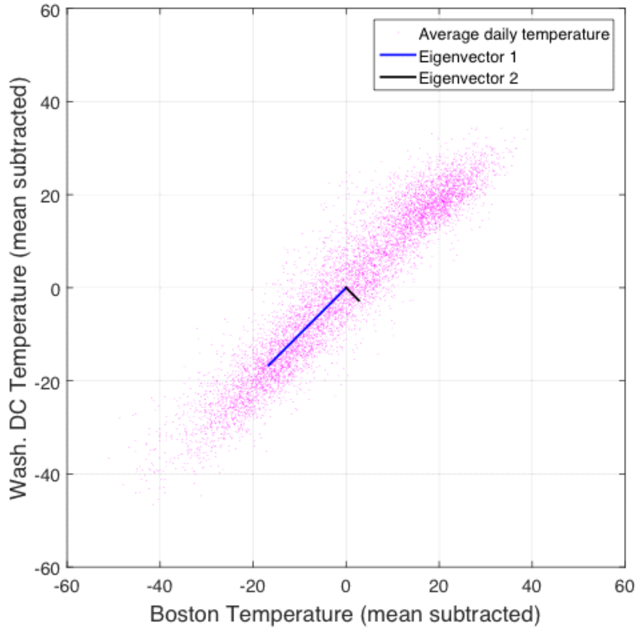


Figure 17.2: Centered average daily temperatures of Boston vs Washington DC, with the eigenvectors of the covariance matrix.

Exercise 17.9

In this next problem, we are going to visualize how the eigenvectors of covariance matrices can tell us about the directions of most variation in 3D data. Load the file `temps_bos_sp_dc.mat` in MATLAB (this file can also be downloaded from the Canvas page for Homework 6). This file will load 21 years of temperature values for Boston, Sao Paolo and Washington DC. Treat the temperatures of Boston, Sao Paolo, and Washington DC for a given day as a point in a 3D space.

1. Subtract out the mean temperature of each city from the daily temperature data.
2. Make a 3D scatter plot of the data points with the means subtracted out. You will find MATLAB's `plot3` function useful. You may wish to use the `'MarkerSize'` argument for `plot3` with a marker size of 0.1 or less to make the plots clearer.
3. Construct a covariance matrix for the data and compute its eigenvectors.
4. On the same axes, using `quiver3`, or `plot3`, plot the eigenvectors scaled by the square-root of their corresponding eigenvalues. Use `grid on` to draw grid lines on the axes to improve your visualization.
5. Using the rotate 3D button on the figure window, rotate the image around to see how the eigenvectors tell you about the variation in the data.

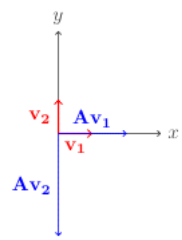
17.5 Diagnostic Quiz

Please see Canvas for the quiz questions.

Solution 17.1

1. First we find
$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 2-\lambda & 0 \\ 0 & -3-\lambda \end{bmatrix}.$$
And then we compute the determinant
$$\det(\mathbf{A} - \lambda\mathbf{I}) = (2-\lambda)(-3-\lambda).$$
Setting this equal to zero produces the characteristic equation,
$$(2-\lambda)(-3-\lambda) = 0$$
whose roots are, in fact, $\lambda_1 = 2$ and $\lambda_2 = -3$.

2. Here's a plot of $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{A}\mathbf{v}_1$ and $\mathbf{A}\mathbf{v}_2$:



When the eigenvalue is negative, the eigenvector is reversed in direction and then scaled.

Solution 17.2

1. First we compute the left-hand side
$$\mathbf{A}\mathbf{v}_1 = \begin{bmatrix} 18 & -2 \\ 12 & 7 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 40 \end{bmatrix}$$
and the right-hand side
$$\lambda_1 \mathbf{v}_1 = 10 \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ 40 \end{bmatrix}.$$
Fortunately, they are equal.

Solution 17.3

Using the fact that $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, we see that
$$\mathbf{A}(c\mathbf{v}) = c\mathbf{A}\mathbf{v} = c\lambda\mathbf{v} = \lambda(c\mathbf{v})$$
and therefore $c\mathbf{v}$ is also an eigenvector.

Solution 17.4

1. We normalize $\mathbf{v}_1$ by first finding its length and then constructing a unit vector. The length of $\mathbf{v}_1$ is given by
$$\|\mathbf{v}_1\| = \sqrt{1^2 + 4^2} = \sqrt{17}$$

Now we construct the unit vector $\hat{\mathbf{v}}_1$ by dividing $\mathbf{v}_1$ by its length

$$\hat{\mathbf{v}}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{bmatrix} 1/\sqrt{17} \\ 4/\sqrt{17} \end{bmatrix}$$

We often use the ˆ symbol to denote a vector that has unit length! Also, notice that even this eigenvector is not unique because we could multiply it by -1 to get the vector $\begin{bmatrix} -1/\sqrt{17} \\ -4/\sqrt{17} \end{bmatrix}$ - all we've done is flip the vector by 180 degrees.

Solution 17.5

First we compute

$$\mathbf{A} - \lambda_2 \mathbf{I} = \begin{bmatrix} 18-15 & -2 \\ 12 & 7-15 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 12 & -8 \end{bmatrix}.$$

Now, letting $\mathbf{v}_2 = \begin{bmatrix} a \\ b \end{bmatrix}$, we are trying to solve

$$\begin{bmatrix} 3 & -2 \\ 12 & -8 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which produces the equations

$$3a - 2b = 0 \text{ and } 12a - 8b = 0.$$

(These equations give the same information since the rows of $(\mathbf{A} - \lambda_2 \mathbf{I})$ are linearly dependent.) This gives $b = \frac{3}{2}a$, so $\mathbf{v}_2 = \begin{bmatrix} a \\ \frac{3}{2}a \end{bmatrix}$ for any value of a. Picking $a = 2$, we have $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. To normalize it we need it's length which is

$$\|\mathbf{v}_2\| = \sqrt{2^2 + 3^2} = \sqrt{13}$$

so that the unit eigenvector is

$$\hat{\mathbf{v}}_2 = \begin{bmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}$$

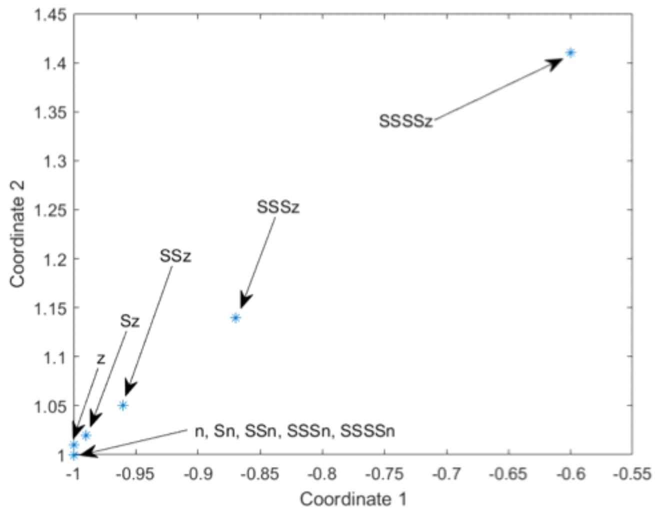
Solution 17.6

1.
$$\lambda_1 = 5, \lambda_2 = 2$$
and
$$\mathbf{v}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

2.
$$\lambda_1 = -1, \lambda_2 = 4$$
and
$$\mathbf{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{bmatrix}$$

Solution 17.7

See Figure 17.1. $\mathbf{n}$ is an eigenvector of $\mathbf{S}$ with an eigenvalue of 1, so it is unchanged by the transformation $\mathbf{S}$. However, $\mathbf{z}$ is not an eigenvector of $\mathbf{S}$, so it changes each time the transformation $\mathbf{S}$ is applied, and the change accelerates as it diverges from the eigenvector $\mathbf{n}$.



Solution 17.8

1.
```
>> A = [18 -2;12 7]

A =

    18    -2
    12     7

>> [V,D] = eig(A)

V =

    0.5547    0.2425
    0.8321    0.9701


D =

    15     0
     0    10
```
You will notice that the eigenvectors are numerical approximations to the exact ones found earlier.

2.
```
>> A = [4 2;1 3]

A =

     4     2
     1     3
```

```
>> [V,D] = eig(A)

V =

    0.8944   -0.7071
    0.4472    0.7071


D =

     5     0
     0     2
```

Great!

3.
```
>> A = [1 2;3 2]

A =

     1     2
     3     2

>> [V,D] = eig(A)

V =

   -0.7071   -0.5547
    0.7071   -0.8321


D =

    -1     0
     0     4
```

You should notice that the eigenvector corresponding to the eigenvalue of 4 has two negative entries instead of two positive ones - this is absolutely fine - it has been "flipped" by 180 degrees. You should not expect the result from MATLAB to match all of your signs, but they should match all of the relative signs!

Solution 17.9

1. `bn=b-mean(b);sn=s-mean(s);wn=w-mean(w);`
2.
```
plot3(bn,sn,wn,'.','MarkerSize',0.1)
xlabel('Boston temperature (mean subtracted)')
ylabel('Sao Paolo temperature (mean subtracted)')
zlabel('Wash. D.C. temperature (mean subtracted)'
```
3.
```
A=1/sqrt(length(b)-1)*[bn,sn,wn];
R=transpose(A)*A;
[V,D]=eig(R)
```
4.
```
plot3(bn,sn,wn,'.','MarkerSize',0.1)
Vs=V.*sqrt(diag(D))
```

```
hold on
plot3([0,Vs(1,1)],[0 Vs(2,1)],[0 Vs(3,1)],'LineWidth',2)
plot3([0,Vs(1,2)],[0 Vs(2,2)],[0 Vs(3,2)],'LineWidth',2)
plot3([0,Vs(1,3)],[0 Vs(2,3)],[0 Vs(3,3)],'LineWidth',2)
grid on
axis equal
```

See Figure 17.3, which has the first eigenvector clearly aligned with the direction of greatest variation.

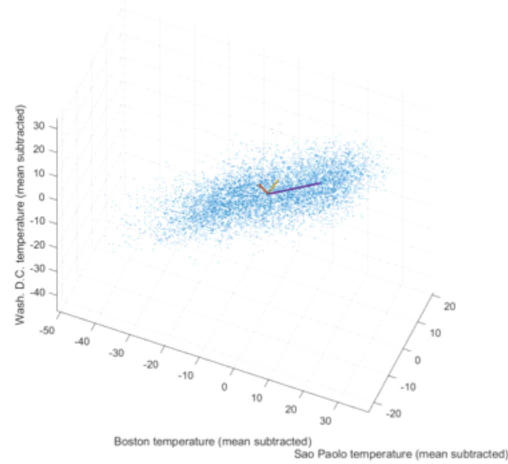


Figure 17.3: Temperatures and eigenvectors.

Untitled

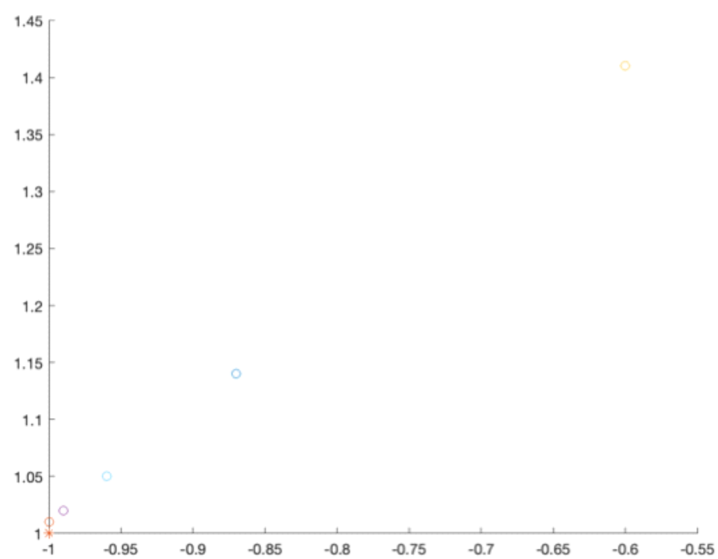
```
oooooo% Exercise 17.7
% I would like to hereby formally recognize my protent against this naming
% scheme.

n = [-1; 1];
z = [-1; 1.01];

S = [2 1; 1 2];

clf; hold on;
plotvec(n, '*');
plotvec(z, 'o');

curN = n;
curZ = z;
for i=1:4
    curN = S * curN;
    curZ = S * curZ;
    plotvec(curN, '*');
    plotvec(curZ, 'o');
end
```




```
% Excercise 17.8
[V, D] = eig([18 -2; 12 7])
```

```
V = 2×2
```

```
    0.5547    0.2425
    0.8321    0.9701
D = 2×2
    15     0
     0    10
```

```
[V, D] = eig([4 2; 1 3])
```

```
V = 2×2
    0.8944   -0.7071
    0.4472    0.7071
D = 2×2
     5     0
     0     2
```

```
[V, D] = eig([1 2; 3 2])
```

```
V = 2×2
   -0.7071   -0.5547
    0.7071   -0.8321
D = 2×2
    -1     0
     0     4
```

```
load('temps_bos_sp_dc.mat');
```

```
% Excercise 17.9
clf; hold on; grid;
dat = [centermean(b), centermean(w), centermean(s)] .* (1 / sqrt(length(b) - 1))
```

```
dat = 7670×3
   -0.1515   -0.2037    0.0379
   -0.1663   -0.2128    0.0574
   -0.2520   -0.3327    0.0482
   -0.2417   -0.2905    0.0482
   -0.3559   -0.4286    0.0779
   -0.2862   -0.3590    0.1133
   -0.0898   -0.2516    0.1224
   -0.2189   -0.2425    0.0711
   -0.2428   -0.2573    0.0802
   -0.3068   -0.2665    0.1122
```

```
covariance = dat' * dat
```

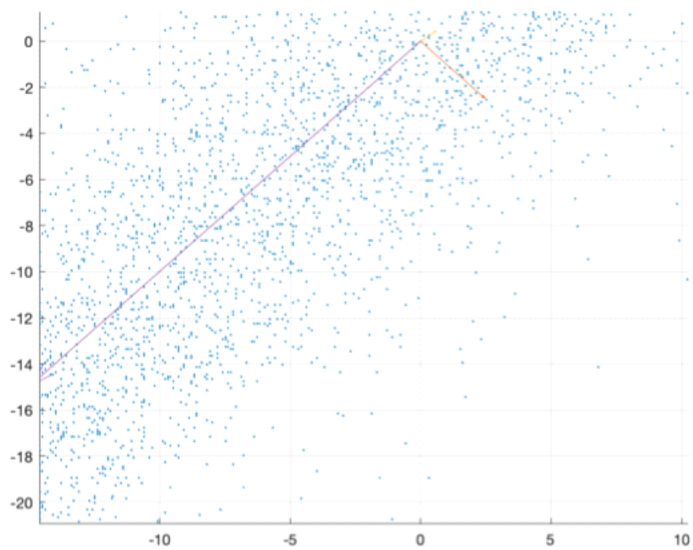
```
covariance = 3×3
  284.8910  269.3927  -56.9729
  269.3927  284.1612  -57.2850
  -56.9729  -57.2850   39.4645
```

```
[V, D] = eig(covariance)
```

```
V = 3×3
```

```
    0.7041    0.1233   -0.6994
   -0.7098    0.0910   -0.6985
   -0.0225    0.9882    0.1516
D = 3×3
   15.1269         0         0
         0   27.0811         0
         0         0  566.3086
```

```
plot3(centermean(b), centermean(w), centermean(s), '*', 'MarkerSize', 0.1)
plotquiver3(V(:, 1) * sqrt(D(1, 1)))
plotquiver3(V(:, 2) * sqrt(D(2, 2)))
plotquiver3(V(:, 3) * sqrt(D(3, 3)))
```




```
function plotquiver3(v)
    quiver3(0, 0, 0, v(1), v(2), v(3))
end

function res = centermean(d)
    res = d - mean(d);
end
function plotvec(v, s)
    plot(v(1), v(2), s)
end
```