

# Chapter 6

## Eigenvalues and Eigenvectors

### 6.1 Introduction to Eigenvalues

- 1 An **eigenvector**  $x$  lies along the same line as  $Ax$  :  $Ax = \lambda x$ . The **eigenvalue** is  $\lambda$ .
- 2 If  $Ax = \lambda x$  then  $A^2x = \lambda^2x$  and  $A^{-1}x = \lambda^{-1}x$  and  $(A + cI)x = (\lambda + c)x$ : the same  $x$ .
- 3 If  $Ax = \lambda x$  then  $(A - \lambda I)x = 0$  and  $A - \lambda I$  is singular and  $\det(A - \lambda I) = 0$ .  $n$  eigenvalues.
- 4 Check  $\lambda$ 's by  $\det A = (\lambda_1)(\lambda_2) \cdots (\lambda_n)$  and diagonal sum  $a_{11} + a_{22} + \cdots + a_{nn}$  = sum of  $\lambda$ 's.
- 5 Projections have  $\lambda = 1$  and 0. Reflections have 1 and  $-1$ . Rotations have  $e^{i\theta}$  and  $e^{-i\theta}$ : *complex!*

This chapter enters a new part of linear algebra. The first part was about  $Ax = b$ : balance and equilibrium and steady state. Now the second part is about **change**. Time enters the picture—continuous time in a differential equation  $du/dt = Au$  or time steps in a difference equation  $u_{k+1} = Au_k$ . Those equations are NOT solved by elimination.

The key idea is to avoid all the complications presented by the matrix  $A$ . Suppose the solution vector  $u(t)$  stays in the direction of a fixed vector  $x$ . Then we only need to find the number (changing with time) that multiplies  $x$ . A number is easier than a vector. **We want “eigenvectors”  $x$  that don’t change direction when you multiply by  $A$ .**

A good model comes from the powers  $A, A^2, A^3, \dots$  of a matrix. Suppose you need the hundredth power  $A^{100}$ . Its columns are very close to the *eigenvector*  $(.6, .4)$ :

$$A, A^2, A^3 = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \quad \begin{bmatrix} .70 & .45 \\ .30 & .55 \end{bmatrix} \quad \begin{bmatrix} .650 & .525 \\ .350 & .475 \end{bmatrix} \quad A^{100} \approx \begin{bmatrix} .6000 & .6000 \\ .4000 & .4000 \end{bmatrix}$$

$A^{100}$  was found by using the *eigenvalues* of  $A$ , not by multiplying 100 matrices. Those eigenvalues (here they are  $\lambda = 1$  and  $1/2$ ) are a new way to see into the heart of a matrix.

To explain eigenvalues, we first explain eigenvectors. Almost all vectors change direction, when they are multiplied by  $A$ . ***Certain exceptional vectors  $x$  are in the same direction as  $Ax$ . Those are the “eigenvectors”.*** Multiply an eigenvector by  $A$ , and the vector  $Ax$  is a number  $\lambda$  times the original  $x$ .

**The basic equation is  $Ax = \lambda x$ . The number  $\lambda$  is an eigenvalue of  $A$ .**

The eigenvalue  $\lambda$  tells whether the special vector  $x$  is stretched or shrunk or reversed or left unchanged—when it is multiplied by  $A$ . We may find  $\lambda = 2$  or  $\frac{1}{2}$  or  $-1$  or  $1$ . The eigenvalue  $\lambda$  could be zero! Then  $Ax = 0x$  means that this eigenvector  $x$  is in the nullspace.

If  $A$  is the identity matrix, every vector has  $Ax = x$ . All vectors are eigenvectors of  $I$ . All eigenvalues “lambda” are  $\lambda = 1$ . This is unusual to say the least. Most 2 by 2 matrices have *two* eigenvector directions and *two* eigenvalues. We will show that  $\det(A - \lambda I) = 0$ .

This section will explain how to compute the  $x$ ’s and  $\lambda$ ’s. It can come early in the course because we only need the determinant of a 2 by 2 matrix. Let me use  $\det(A - \lambda I) = 0$  to find the eigenvalues for this first example, and then derive it properly in equation (3).

**Example 1** The matrix  $A$  has two eigenvalues  $\lambda = 1$  and  $\lambda = 1/2$ . Look at  $\det(A - \lambda I)$ :

$$A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \quad \det \begin{bmatrix} .8 - \lambda & .3 \\ .2 & .7 - \lambda \end{bmatrix} = \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = (\lambda - 1) \left( \lambda - \frac{1}{2} \right).$$

I factored the quadratic into  $\lambda - 1$  times  $\lambda - \frac{1}{2}$ , to see the two eigenvalues  $\lambda = 1$  and  $\lambda = \frac{1}{2}$ . For those numbers, the matrix  $A - \lambda I$  becomes *singular* (zero determinant). The eigenvectors  $x_1$  and  $x_2$  are in the nullspaces of  $A - I$  and  $A - \frac{1}{2}I$ .

$(A - I)x_1 = 0$  is  $Ax_1 = x_1$  and the first eigenvector is  $(.6, .4)$ .

$(A - \frac{1}{2}I)x_2 = 0$  is  $Ax_2 = \frac{1}{2}x_2$  and the second eigenvector is  $(1, -1)$ :

$$\begin{aligned} x_1 &= \begin{bmatrix} .6 \\ .4 \end{bmatrix} \quad \text{and} \quad Ax_1 = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \begin{bmatrix} .6 \\ .4 \end{bmatrix} = x_1 \quad (Ax = x \text{ means that } \lambda_1 = 1) \\ x_2 &= \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad Ax_2 = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} .5 \\ -.5 \end{bmatrix} \quad (\text{this is } \frac{1}{2}x_2 \text{ so } \lambda_2 = \frac{1}{2}). \end{aligned}$$

If  $x_1$  is multiplied again by  $A$ , we still get  $x_1$ . Every power of  $A$  will give  $A^n x_1 = x_1$ . Multiplying  $x_2$  by  $A$  gave  $\frac{1}{2}x_2$ , and if we multiply again we get  $(\frac{1}{2})^2$  times  $x_2$ .

***When  $A$  is squared, the eigenvectors stay the same. The eigenvalues are squared.***

This pattern keeps going, because the eigenvectors stay in their own directions (Figure 6.1) and never get mixed. The eigenvectors of  $A^{100}$  are the same  $x_1$  and  $x_2$ . The eigenvalues of  $A^{100}$  are  $1^{100} = 1$  and  $(\frac{1}{2})^{100}$  = very small number.

Other vectors do change direction. But all other vectors are combinations of the two eigenvectors. The first column of  $A$  is the combination  $x_1 + (.2)x_2$ :

$$\begin{array}{l} \text{Separate into eigenvectors} \\ \text{Then multiply by } A \end{array} \quad \begin{bmatrix} .8 \\ .2 \end{bmatrix} = x_1 + (.2)x_2 = \begin{bmatrix} .6 \\ .4 \end{bmatrix} + \begin{bmatrix} .2 \\ -.2 \end{bmatrix}. \quad (1)$$

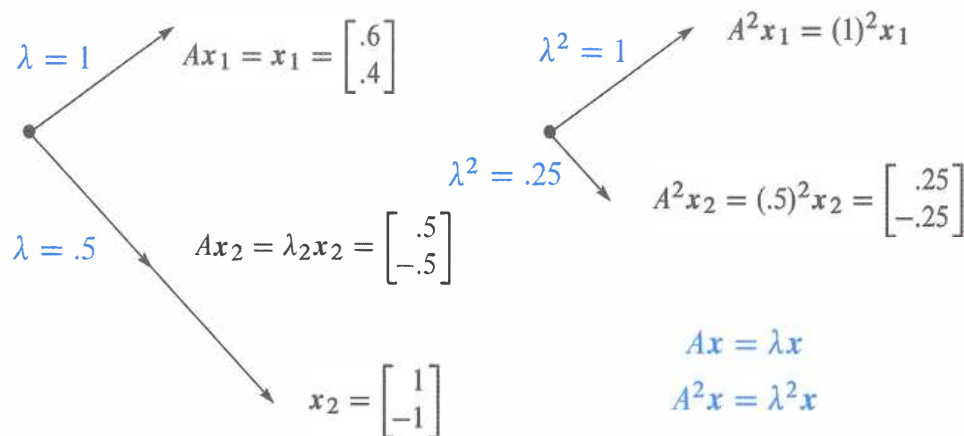


Figure 6.1: The eigenvectors keep their directions.  $A^2x = \lambda^2x$  with  $\lambda^2 = 1^2$  and  $(.5)^2$ .

When we multiply separately for  $x_1$  and  $(.2)x_2$ ,  $A$  multiplies  $x_2$  by its eigenvalue  $\frac{1}{2}$ :

**Multiply each  $x_i$  by  $\lambda_i$**      $A \begin{bmatrix} .8 \\ .2 \end{bmatrix}$     is     $x_1 + \frac{1}{2}(.2)x_2 = \begin{bmatrix} .6 \\ .4 \end{bmatrix} + \begin{bmatrix} .1 \\ -.1 \end{bmatrix} = \begin{bmatrix} .7 \\ .3 \end{bmatrix}.$

*Each eigenvector is multiplied by its eigenvalue*, when we multiply by  $A$ . At every step  $x_1$  is unchanged and  $x_2$  is multiplied by  $(\frac{1}{2})$ , so 99 steps give the small number  $(\frac{1}{2})^{99}$ :

$$A^{99} \begin{bmatrix} .8 \\ .2 \end{bmatrix} \text{ is really } x_1 + (.2)(\frac{1}{2})^{99} x_2 = \begin{bmatrix} .6 \\ .4 \end{bmatrix} + \begin{bmatrix} \text{very} \\ \text{small} \\ \text{vector} \end{bmatrix}.$$

This is the first column of  $A^{100}$ . The number we originally wrote as .6000 was not exact. We left out  $(.2)(\frac{1}{2})^{99}$  which wouldn't show up for 30 decimal places.

The eigenvector  $x_1$  is a “steady state” that doesn't change (because  $\lambda_1 = 1$ ). The eigenvector  $x_2$  is a “decaying mode” that virtually disappears (because  $\lambda_2 = .5$ ). The higher the power of  $A$ , the more closely its columns approach the steady state.

This particular  $A$  is a **Markov matrix**. Its largest eigenvalue is  $\lambda = 1$ . Its eigenvector  $x_1 = (.6, .4)$  is the *steady state*—which all columns of  $A^k$  will approach. Section 10.3 shows how Markov matrices appear when you search with Google.

*For projection matrices  $P$ , we can see when  $Px$  is parallel to  $x$ .* The eigenvectors for  $\lambda = 1$  and  $\lambda = 0$  fill the column space and nullspace. The column space doesn't move ( $Px = x$ ). The nullspace goes to zero ( $Px = 0x$ ).

**Example 2**

The projection matrix  $P = \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}$  has eigenvalues  $\lambda = 1$  and  $\lambda = 0$ .

Its eigenvectors are  $x_1 = (1, 1)$  and  $x_2 = (1, -1)$ . For those vectors,  $Px_1 = x_1$  (steady state) and  $Px_2 = \mathbf{0}$  (nullspace). This example illustrates Markov matrices and singular matrices and (most important) symmetric matrices. All have special  $\lambda$ 's and  $x$ 's:

1. **Markov matrix**: Each column of  $P$  adds to 1, so  $\lambda = 1$  is an eigenvalue.
2.  $P$  is **singular**, so  $\lambda = 0$  is an eigenvalue.
3.  $P$  is **symmetric**, so its eigenvectors  $(1, 1)$  and  $(1, -1)$  are perpendicular.

The only eigenvalues of a projection matrix are 0 and 1. The eigenvectors for  $\lambda = 0$  (which means  $Px = 0x$ ) fill up the nullspace. The eigenvectors for  $\lambda = 1$  (which means  $Px = x$ ) fill up the column space. The nullspace is projected to zero. The column space projects onto itself. The projection keeps the column space and destroys the nullspace:

**Project each part**  $v = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \end{bmatrix}$  **projects onto**  $Pv = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ .

Projections have  $\lambda = 0$  and 1. Permutations have all  $|\lambda| = 1$ . The next matrix  $R$  is a reflection and at the same time a permutation.  $R$  also has special eigenvalues.

**Example 3**

The reflection matrix  $R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  has eigenvalues 1 and  $-1$ .

The eigenvector  $(1, 1)$  is unchanged by  $R$ . The second eigenvector is  $(1, -1)$ —its signs are reversed by  $R$ . A matrix with no negative entries can still have a negative eigenvalue! The eigenvectors for  $R$  are the same as for  $P$ , because *reflection* = *2(projection)*  $- I$ :

$$R = 2P - I \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 2 \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (2)$$

*When a matrix is shifted by  $I$ , each  $\lambda$  is shifted by 1.* No change in eigenvectors.

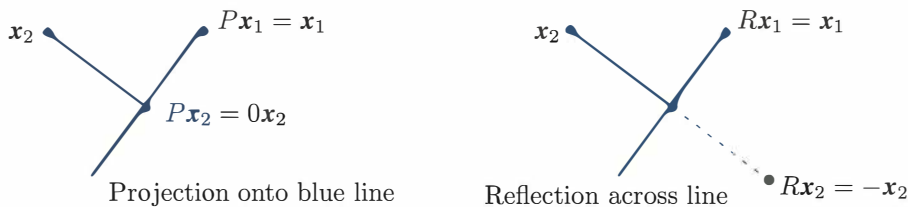


Figure 6.2: Projections  $P$  have eigenvalues 1 and 0. Reflections  $R$  have  $\lambda = 1$  and  $-1$ . A typical  $x$  changes direction, but an eigenvector stays along the same line.

### The Equation for the Eigenvalues

For projection matrices we found  $\lambda$ 's and  $x$ 's by geometry:  $Px = x$  and  $Px = 0$ . For other matrices we use determinants and linear algebra. *This is the key calculation in the chapter*—almost every application starts by solving  $Ax = \lambda x$ .

**First move  $\lambda x$  to the left side.** Write the equation  $Ax = \lambda x$  as  $(A - \lambda I)x = 0$ . The matrix  $A - \lambda I$  times the eigenvector  $x$  is the zero vector. **The eigenvectors make up the nullspace of  $A - \lambda I$ .** When we know an eigenvalue  $\lambda$ , we find an eigenvector by solving  $(A - \lambda I)x = 0$ .

Eigenvalues first. If  $(A - \lambda I)x = 0$  has a nonzero solution,  $A - \lambda I$  is not invertible. **The determinant of  $A - \lambda I$  must be zero.** This is how to recognize an eigenvalue  $\lambda$ :

**Eigenvalues** The number  $\lambda$  is an eigenvalue of  $A$  if and only if  $A - \lambda I$  is singular.

**Equation for the eigenvalues**  $\det(A - \lambda I) = 0.$  (3)

This “characteristic polynomial”  $\det(A - \lambda I)$  involves only  $\lambda$ , not  $x$ . When  $A$  is  $n$  by  $n$ , equation (3) has degree  $n$ . Then  $A$  has  $n$  eigenvalues (repeats possible!) Each  $\lambda$  leads to  $x$ :

**For each eigenvalue  $\lambda$  solve  $(A - \lambda I)x = 0$  or  $Ax = \lambda x$  to find an eigenvector  $x$ .**

**Example 4**  $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$  is already singular (zero determinant). Find its  $\lambda$ 's and  $x$ 's.

When  $A$  is singular,  $\lambda = 0$  is one of the eigenvalues. The equation  $Ax = 0x$  has solutions. They are the eigenvectors for  $\lambda = 0$ . But  $\det(A - \lambda I) = 0$  is the way to find *all*  $\lambda$ 's and  $x$ 's. Always subtract  $\lambda I$  from  $A$ :

**Subtract  $\lambda$  from the diagonal to find**  $A - \lambda I = \begin{bmatrix} 1 - \lambda & 2 \\ 2 & 4 - \lambda \end{bmatrix}.$  (4)

**Take the determinant “ $ad - bc$ ” of this 2 by 2 matrix.** From  $1 - \lambda$  times  $4 - \lambda$ , the “ $ad$ ” part is  $\lambda^2 - 5\lambda + 4$ . The “ $bc$ ” part, not containing  $\lambda$ , is 2 times 2.

$$\det \begin{bmatrix} 1 - \lambda & 2 \\ 2 & 4 - \lambda \end{bmatrix} = (1 - \lambda)(4 - \lambda) - (2)(2) = \lambda^2 - 5\lambda. \quad (5)$$

**Set this determinant  $\lambda^2 - 5\lambda$  to zero.** One solution is  $\lambda = 0$  (as expected, since  $A$  is singular). Factoring into  $\lambda$  times  $\lambda - 5$ , the other root is  $\lambda = 5$ :

$\det(A - \lambda I) = \lambda^2 - 5\lambda = 0$  yields the eigenvalues  $\lambda_1 = 0$  and  $\lambda_2 = 5$ .

Now find the eigenvectors. Solve  $(A - \lambda I)x = 0$  separately for  $\lambda_1 = 0$  and  $\lambda_2 = 5$ :

$$(A - 0I)x = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ yields an eigenvector } \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \text{ for } \lambda_1 = 0$$

$$(A - 5I)x = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ yields an eigenvector } \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ for } \lambda_2 = 5.$$

The matrices  $A - 0I$  and  $A - 5I$  are singular (because 0 and 5 are eigenvalues). The eigenvectors  $(2, -1)$  and  $(1, 2)$  are in the nullspaces:  $(A - \lambda I)x = 0$  is  $Ax = \lambda x$ .

We need to emphasize: *There is nothing exceptional about  $\lambda = 0$ .* Like every other number, zero might be an eigenvalue and it might not. If  $A$  is singular, the eigenvectors for  $\lambda = 0$  fill the nullspace:  $Ax = 0x = 0$ . If  $A$  is invertible, zero is not an eigenvalue. We shift  $A$  by a multiple of  $I$  to make it singular.

In the example, the shifted matrix  $A - 5I$  is singular and 5 is the other eigenvalue.

**Summary** To solve the eigenvalue problem for an  $n$  by  $n$  matrix, follow these steps:

1. **Compute the determinant of  $A - \lambda I$ .** With  $\lambda$  subtracted along the diagonal, this determinant starts with  $\lambda^n$  or  $-\lambda^n$ . It is a polynomial in  $\lambda$  of degree  $n$ .
2. **Find the roots of this polynomial,** by solving  $\det(A - \lambda I) = 0$ . The  $n$  roots are the  $n$  eigenvalues of  $A$ . They make  $A - \lambda I$  singular.
3. For each eigenvalue  $\lambda$ , **solve  $(A - \lambda I)x = 0$  to find an eigenvector  $x$ .**

A note on the eigenvectors of 2 by 2 matrices. When  $A - \lambda I$  is singular, both rows are multiples of a vector  $(a, b)$ . *The eigenvector is any multiple of  $(b, -a)$ .* The example had

$\lambda = 0$  : rows of  $A - 0I$  in the direction  $(1, 2)$ ; eigenvector in the direction  $(2, -1)$

$\lambda = 5$  : rows of  $A - 5I$  in the direction  $(-4, 2)$ ; eigenvector in the direction  $(2, 4)$ .

Previously we wrote that last eigenvector as  $(1, 2)$ . Both  $(1, 2)$  and  $(2, 4)$  are correct. There is a whole *line of eigenvectors*—any nonzero multiple of  $x$  is as good as  $x$ . MATLAB's `eig(A)` divides by the length, to make the eigenvector into a unit vector.

We must add a warning. Some 2 by 2 matrices have only *one* line of eigenvectors. This can only happen when two eigenvalues are equal. (On the other hand  $A = I$  has equal eigenvalues and plenty of eigenvectors.) Without a full set of eigenvectors, we don't have a basis. We can't write every  $v$  as a combination of eigenvectors. In the language of the next section, *we can't diagonalize a matrix without  $n$  independent eigenvectors.*

## Determinant and Trace

Bad news first: If you add a row of  $A$  to another row, or exchange rows, the eigenvalues usually change. *Elimination does not preserve the  $\lambda$ 's.* The triangular  $U$  has its eigenvalues sitting along the diagonal—they are the pivots. But they are not the eigenvalues of  $A$ ! Eigenvalues are changed when row 1 is added to row 2:

$$U = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix} \quad \text{has } \lambda = 0 \text{ and } \lambda = 1; \quad A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \quad \text{has } \lambda = 0 \text{ and } \lambda = 7.$$

Good news second: The product  $\lambda_1$  times  $\lambda_2$  and the sum  $\lambda_1 + \lambda_2$  can be found quickly from the matrix. For this  $A$ , the product is 0 times 7. That agrees with the determinant (which is 0). The sum of eigenvalues is 0 + 7. That agrees with the sum down the main diagonal (the **trace** is 1 + 6). These quick checks always work:

*The product of the  $n$  eigenvalues equals the determinant.  
The sum of the  $n$  eigenvalues equals the sum of the  $n$  diagonal entries.*

The sum of the entries along the main diagonal is called the **trace** of  $A$ :

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \text{trace} = a_{11} + a_{22} + \cdots + a_{nn}. \quad (6)$$

Those checks are very useful. They are proved in Problems 16–17 and again in the next section. They don't remove the pain of computing  $\lambda$ 's. But when the computation is wrong, they generally tell us so. To compute the correct  $\lambda$ 's, go back to  $\det(A - \lambda I) = 0$ .

The trace and determinant *do* tell everything when the matrix is 2 by 2. We never want to get those wrong! Here trace = 3 and det = 2, so the eigenvalues are  $\lambda = 1$  and  $\lambda = 2$ :

$$A = \begin{bmatrix} 1 & 9 \\ 0 & 2 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 7 & -3 \\ 10 & -4 \end{bmatrix}. \quad (7)$$

And here is a question about the best matrices for finding eigenvalues: *triangular*.

**Why do the eigenvalues of a triangular matrix lie along its diagonal?**

## Imaginary Eigenvalues

One more bit of news (not too terrible). The eigenvalues might not be real numbers.

**Example 5** The 90° rotation  $Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$  has no real eigenvectors. Its eigenvalues are  $\lambda_1 = i$  and  $\lambda_2 = -i$ . Then  $\lambda_1 + \lambda_2 = \text{trace} = 0$  and  $\lambda_1 \lambda_2 = \text{determinant} = 1$ .

After a rotation, no real vector  $Qx$  stays in the same direction as  $x$  ( $x = 0$  is useless). There cannot be an eigenvector, unless we go to **imaginary numbers**. Which we do.

To see how  $i = \sqrt{-1}$  can help, look at  $Q^2$  which is  $-I$ . If  $Q$  is rotation through  $90^\circ$ , then  $Q^2$  is rotation through  $180^\circ$ . Its eigenvalues are  $-1$  and  $-1$ . (Certainly  $-Ix = -1x$ .) Squaring  $Q$  will square each  $\lambda$ , so we must have  $\lambda^2 = -1$ . *The eigenvalues of the  $90^\circ$  rotation matrix  $Q$  are  $+i$  and  $-i$ , because  $i^2 = -1$ .*

Those  $\lambda$ 's come as usual from  $\det(Q - \lambda I) = 0$ . This equation gives  $\lambda^2 + 1 = 0$ . Its roots are  $i$  and  $-i$ . We meet the imaginary number  $i$  also in the eigenvectors:

$$\text{Complex eigenvectors} \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = -i \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = i \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

Somehow these complex vectors  $x_1 = (1, i)$  and  $x_2 = (i, 1)$  keep their direction as they are rotated. Don't ask me how. This example makes the all-important point that real matrices can easily have complex eigenvalues and eigenvectors. The particular eigenvalues  $i$  and  $-i$  also illustrate two special properties of  $Q$ :

1.  $Q$  is an orthogonal matrix so the absolute value of each  $\lambda$  is  $|\lambda| = 1$ .
2.  $Q$  is a skew-symmetric matrix so each  $\lambda$  is pure imaginary.

A symmetric matrix ( $S^T = S$ ) can be compared to a real number. A skew-symmetric matrix ( $A^T = -A$ ) can be compared to an imaginary number. An orthogonal matrix ( $Q^T Q = I$ ) corresponds to a complex number with  $|\lambda| = 1$ . For the eigenvalues of  $S$  and  $A$  and  $Q$ , those are more than analogies—they are facts to be proved in Section 6.4.

The eigenvectors for all these special matrices are perpendicular. Somehow  $(i, 1)$  and  $(1, i)$  are perpendicular (Chapter 9 explains the dot product of complex vectors).

### Eigenvalues of $AB$ and $A+B$

The first guess about the eigenvalues of  $AB$  is not true. An eigenvalue  $\lambda$  of  $A$  times an eigenvalue  $\beta$  of  $B$  usually does *not* give an eigenvalue of  $AB$ :

$$\text{False proof} \quad ABx = A\beta x = \beta Ax = \beta\lambda x. \quad (8)$$

It seems that  $\beta$  times  $\lambda$  is an eigenvalue. When  $x$  is an eigenvector for  $A$  and  $B$ , this proof is correct. ***The mistake is to expect that  $A$  and  $B$  automatically share the same eigenvector  $x$ .*** Usually they don't. Eigenvectors of  $A$  are not generally eigenvectors of  $B$ .  $A$  and  $B$  could have all zero eigenvalues while 1 is an eigenvalue of  $AB$ :

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}; \quad \text{then} \quad AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A+B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

For the same reason, the eigenvalues of  $A+B$  are generally not  $\lambda + \beta$ . Here  $\lambda + \beta = 0$  while  $A+B$  has eigenvalues 1 and  $-1$ . (At least they add to zero.)

The false proof suggests what is true. Suppose  $x$  really is an eigenvector for both  $A$  and  $B$ . Then we do have  $ABx = \lambda\beta x$  and  $BAx = \lambda\beta x$ . When all  $n$  eigenvectors are shared, we *can* multiply eigenvalues. The test  $AB = BA$  for shared eigenvectors is important in quantum mechanics—time out to mention this application of linear algebra:

$A$  and  $B$  share the same  $n$  independent eigenvectors if and only if  $AB = BA$ .

**Heisenberg's uncertainty principle** In quantum mechanics, the position matrix  $P$  and the momentum matrix  $Q$  do not commute. In fact  $QP - PQ = I$  (these are infinite matrices). To have  $Px = 0$  at the same time as  $Qx = 0$  would require  $x = Ix = 0$ . If we knew the position exactly, we could not also know the momentum exactly. Problem 36 derives Heisenberg's uncertainty principle  $\|Px\| \|Qx\| \geq \frac{1}{2} \|x\|^2$ .

## ■ REVIEW OF THE KEY IDEAS ■

1.  $Ax = \lambda x$  says that eigenvectors  $x$  keep the same direction when multiplied by  $A$ .
2.  $Ax = \lambda x$  also says that  $\det(A - \lambda I) = 0$ . This determines  $n$  eigenvalues.
3. The eigenvalues of  $A^2$  and  $A^{-1}$  are  $\lambda^2$  and  $\lambda^{-1}$ , with the same eigenvectors.
4. The sum of the  $\lambda$ 's equals the sum down the main diagonal of  $A$  (*the trace*). The product of the  $\lambda$ 's equals the determinant of  $A$ .
5. Projections  $P$ , reflections  $R$ ,  $90^\circ$  rotations  $Q$  have special eigenvalues  $1, 0, -1, i, -i$ . Singular matrices have  $\lambda = 0$ . Triangular matrices have  $\lambda$ 's on their diagonal.
6. *Special properties of a matrix lead to special eigenvalues and eigenvectors.* That is a major theme of this chapter (it is captured in a table at the very end).

## ■ WORKED EXAMPLES ■

**6.1 A** Find the eigenvalues and eigenvectors of  $A$  and  $A^2$  and  $A^{-1}$  and  $A + 4I$ :

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \quad \text{and} \quad A^2 = \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}.$$

Check the trace  $\lambda_1 + \lambda_2 = 4$  and the determinant  $\lambda_1 \lambda_2 = 3$ .

**Solution** The eigenvalues of  $A$  come from  $\det(A - \lambda I) = 0$ :

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \quad \det(A - \lambda I) = \begin{vmatrix} 2 - \lambda & -1 \\ -1 & 2 - \lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3 = 0.$$

This factors into  $(\lambda - 1)(\lambda - 3) = 0$  so the eigenvalues of  $A$  are  $\lambda_1 = 1$  and  $\lambda_2 = 3$ . For the trace, the sum  $2 + 2$  agrees with  $1 + 3$ . The determinant 3 agrees with the product  $\lambda_1 \lambda_2$ .

The eigenvectors come separately by solving  $(A - \lambda I)x = \mathbf{0}$  which is  $Ax = \lambda x$ :

$$\lambda = 1: (A - I)x = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ gives the eigenvector } x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda = 3: (A - 3I)x = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ gives the eigenvector } x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$A^2$  and  $A^{-1}$  and  $A + 4I$  keep the *same eigenvectors* as  $A$ . Their eigenvalues are  $\lambda^2$  and  $\lambda^{-1}$  and  $\lambda + 4$ :

$$A^2 \text{ has eigenvalues } 1^2 = 1 \text{ and } 3^2 = 9 \quad A^{-1} \text{ has } \frac{1}{1} \text{ and } \frac{1}{3} \quad A + 4I \text{ has } \frac{1+4=5}{3+4=7}$$

Notes for later sections:  $A$  has *orthogonal eigenvectors* (Section 6.4 on symmetric matrices).  $A$  can be *diagonalized* since  $\lambda_1 \neq \lambda_2$  (Section 6.2).  $A$  is *similar* to any 2 by 2 matrix with eigenvalues 1 and 3 (Section 6.2).  $A$  is a *positive definite matrix* (Section 6.5) since  $A = A^T$  and the  $\lambda$ 's are positive.

### 6.1 B How can you estimate the eigenvalues of any $A$ ? Gershgorin gave this answer.

Every eigenvalue of  $A$  must be “near” at least one of the entries  $a_{ii}$  on the main diagonal. For  $\lambda$  to be “near  $a_{ii}$ ” means that  $|a_{ii} - \lambda|$  is no more than **the sum  $R_i$  of all other  $|a_{ij}|$  in that row  $i$  of the matrix**. Then  $R_i = \sum_{j \neq i} |a_{ij}|$  is the radius of a circle centered at  $a_{ii}$ .

**Every  $\lambda$  is in the circle around one or more diagonal entries  $a_{ii}$ :  $|a_{ii} - \lambda| \leq R_i$ .**

Here is the reasoning. If  $\lambda$  is an eigenvalue, then  $A - \lambda I$  is not invertible. Then  $A - \lambda I$  cannot be diagonally dominant (see Section 2.5). So at least one diagonal entry  $a_{ii} - \lambda$  is *not larger* than the sum  $R_i$  of all other entries  $|a_{ij}|$  (we take absolute values!) in row  $i$ .

*Example 1.* Every eigenvalue  $\lambda$  of this  $A$  falls into one or both of the **Gershgorin circles**: The centers are  $a$  and  $d$ , the radii are  $R_1 = |b|$  and  $R_2 = |c|$ .

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \begin{array}{ll} \text{First circle:} & |\lambda - a| \leq |b| \\ \text{Second circle:} & |\lambda - d| \leq |c| \end{array}$$

Those are circles in the complex plane, since  $\lambda$  could certainly be complex.

*Example 2.* All eigenvalues of this  $A$  lie in a circle of radius  $R = 3$  around *one or more* of the diagonal entries  $d_1, d_2, d_3$ :

$$A = \begin{bmatrix} d_1 & 1 & 2 \\ 2 & d_2 & 1 \\ -1 & 2 & d_3 \end{bmatrix} \quad \begin{array}{l} |\lambda - d_1| \leq 1 + 2 = R_1 \\ |\lambda - d_2| \leq 2 + 1 = R_2 \\ |\lambda - d_3| \leq 1 + 2 = R_3 \end{array}$$

You see that “near” means not more than 3 away from  $d_1$  or  $d_2$  or  $d_3$ , for this example.

**6.1 C** Find the eigenvalues and eigenvectors of this symmetric 3 by 3 matrix  $S$ :

**Symmetric matrix**

**Singular matrix**

**Trace  $1 + 2 + 1 = 4$**

$$S = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

**Solution** Since all rows of  $S$  add to zero, the vector  $\mathbf{x} = (1, 1, 1)$  gives  $S\mathbf{x} = \mathbf{0}$ . This is an eigenvector for  $\lambda = 0$ . To find  $\lambda_2$  and  $\lambda_3$  I will compute the 3 by 3 determinant:

$$\det(S - \lambda I) = \begin{vmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)(1-\lambda) - 2(1-\lambda) \\ = (1-\lambda)[(2-\lambda)(1-\lambda) - 2] \\ = (1-\lambda)(-\lambda)(3-\lambda).$$

Those three factors give  $\lambda = 0, 1, 3$ . Each eigenvalue corresponds to an eigenvector (or a line of eigenvectors):

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad S\mathbf{x}_1 = \mathbf{0}\mathbf{x}_1 \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad S\mathbf{x}_2 = \mathbf{1}\mathbf{x}_2 \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad S\mathbf{x}_3 = \mathbf{3}\mathbf{x}_3.$$

I notice again that eigenvectors are perpendicular when  $S$  is symmetric. We were lucky to find  $\lambda = 0, 1, 3$ . For a larger matrix I would use **eig**( $A$ ), and never touch determinants.

The full command **[X,E]=eig(A)** will produce unit eigenvectors in the columns of  $X$ .

## Problem Set 6.1

**1** The example at the start of the chapter has powers of this matrix  $A$ :

$$A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \quad \text{and} \quad A^2 = \begin{bmatrix} .70 & .45 \\ .30 & .55 \end{bmatrix} \quad \text{and} \quad A^\infty = \begin{bmatrix} .6 & .6 \\ .4 & .4 \end{bmatrix}.$$

Find the eigenvalues of these matrices. All powers have the same eigenvectors.

- Show from  $A$  how a row exchange can produce different eigenvalues.
- Why is a zero eigenvalue *not* changed by the steps of elimination?

**2** Find the eigenvalues and the eigenvectors of these two matrices:

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \quad \text{and} \quad A + I = \begin{bmatrix} 2 & 4 \\ 2 & 4 \end{bmatrix}.$$

$A + I$  has the \_\_\_\_\_ eigenvectors as  $A$ . Its eigenvalues are \_\_\_\_\_ by 1.

**3** Compute the eigenvalues and eigenvectors of  $A$  and  $A^{-1}$ . Check the trace !

$$A = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad A^{-1} = \begin{bmatrix} -1/2 & 1 \\ 1/2 & 0 \end{bmatrix}.$$

$A^{-1}$  has the \_\_\_\_\_ eigenvectors as  $A$ . When  $A$  has eigenvalues  $\lambda_1$  and  $\lambda_2$ , its inverse has eigenvalues \_\_\_\_\_.

- 4 Compute the eigenvalues and eigenvectors of  $A$  and  $A^2$ :

$$A = \begin{bmatrix} -1 & 3 \\ 2 & 0 \end{bmatrix} \quad \text{and} \quad A^2 = \begin{bmatrix} 7 & -3 \\ -2 & 6 \end{bmatrix}.$$

$A^2$  has the same \_\_\_\_\_ as  $A$ . When  $A$  has eigenvalues  $\lambda_1$  and  $\lambda_2$ ,  $A^2$  has eigenvalues \_\_\_\_\_. In this example, why is  $\lambda_1^2 + \lambda_2^2 = 13$ ?

- 5 Find the eigenvalues of  $A$  and  $B$  (easy for triangular matrices) and  $A + B$ :

$$A = \begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad A + B = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}.$$

Eigenvalues of  $A + B$  (are equal to)(are not equal to) eigenvalues of  $A$  plus eigenvalues of  $B$ .

- 6 Find the eigenvalues of  $A$  and  $B$  and  $AB$  and  $BA$ :

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad AB = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}.$$

- (a) Are the eigenvalues of  $AB$  equal to eigenvalues of  $A$  times eigenvalues of  $B$ ?  
 (b) Are the eigenvalues of  $AB$  equal to the eigenvalues of  $BA$ ?
- 7 Elimination produces  $A = LU$ . The eigenvalues of  $U$  are on its diagonal; they are the \_\_\_\_\_. The eigenvalues of  $L$  are on its diagonal; they are all \_\_\_\_\_. The eigenvalues of  $A$  are not the same as \_\_\_\_\_.
- 8 (a) If you know that  $\mathbf{x}$  is an eigenvector, the way to find  $\lambda$  is to \_\_\_\_\_.  
 (b) If you know that  $\lambda$  is an eigenvalue, the way to find  $\mathbf{x}$  is to \_\_\_\_\_.
- 9 What do you do to the equation  $A\mathbf{x} = \lambda\mathbf{x}$ , in order to prove (a), (b), and (c)?
- (a)  $\lambda^2$  is an eigenvalue of  $A^2$ , as in Problem 4.  
 (b)  $\lambda^{-1}$  is an eigenvalue of  $A^{-1}$ , as in Problem 3.  
 (c)  $\lambda + 1$  is an eigenvalue of  $A + I$ , as in Problem 2.
- 10 Find the eigenvalues and eigenvectors for both of these Markov matrices  $A$  and  $A^\infty$ . Explain from those answers why  $A^{100}$  is close to  $A^\infty$ :

$$A = \begin{bmatrix} .6 & .2 \\ .4 & .8 \end{bmatrix} \quad \text{and} \quad A^\infty = \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & 2/3 \end{bmatrix}.$$

- 11 Here is a strange fact about 2 by 2 matrices with eigenvalues  $\lambda_1 \neq \lambda_2$ : The columns of  $A - \lambda_1 I$  are multiples of the eigenvector  $\mathbf{x}_2$ . Any idea why this should be?

- 12 Find three eigenvectors for this matrix  $P$  (projection matrices have  $\lambda = 1$  and  $0$ ):

$$\text{Projection matrix} \quad P = \begin{bmatrix} .2 & .4 & 0 \\ .4 & .8 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

If two eigenvectors share the same  $\lambda$ , so do all their linear combinations. Find an eigenvector of  $P$  with no zero components.

- 13 From the unit vector  $u = (\frac{1}{6}, \frac{1}{6}, \frac{3}{6}, \frac{5}{6})$  construct the rank one projection matrix  $P = uu^T$ . This matrix has  $P^2 = P$  because  $u^T u = 1$ .

(a)  $Pu = u$  comes from  $(uu^T)u = u(\text{_____})$ . Then  $u$  is an eigenvector with  $\lambda = 1$ .

(b) If  $v$  is perpendicular to  $u$  show that  $Pv = 0$ . Then  $\lambda = 0$ .

(c) Find three independent eigenvectors of  $P$  all with eigenvalue  $\lambda = 0$ .

- 14 Solve  $\det(Q - \lambda I) = 0$  by the quadratic formula to reach  $\lambda = \cos \theta \pm i \sin \theta$ :

$$Q = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad \text{rotates the } xy \text{ plane by the angle } \theta. \text{ No real } \lambda\text{'s.}$$

Find the eigenvectors of  $Q$  by solving  $(Q - \lambda I)x = 0$ . Use  $i^2 = -1$ .

- 15 Every permutation matrix leaves  $x = (1, 1, \dots, 1)$  unchanged. Then  $\lambda = 1$ . Find two more  $\lambda$ 's (possibly complex) for these permutations, from  $\det(P - \lambda I) = 0$ :

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

- 16 **The determinant of  $A$  equals the product  $\lambda_1 \lambda_2 \cdots \lambda_n$ .** Start with the polynomial  $\det(A - \lambda I)$  separated into its  $n$  factors (always possible). Then set  $\lambda = 0$ :

$$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda) \quad \text{so} \quad \det A = \text{_____}.$$

Check this rule in Example 1 where the Markov matrix has  $\lambda = 1$  and  $\frac{1}{2}$ .

- 17 The sum of the diagonal entries (the *trace*) equals the sum of the eigenvalues:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{has} \quad \det(A - \lambda I) = \lambda^2 - (a + d)\lambda + ad - bc = 0.$$

The quadratic formula gives the eigenvalues  $\lambda = (a + d + \sqrt{\text{_____}})/2$  and  $\lambda = \text{_____}$ . Their sum is \_\_\_\_\_. If  $A$  has  $\lambda_1 = 3$  and  $\lambda_2 = 4$  then  $\det(A - \lambda I) = \text{_____}$ .

- 18 If  $A$  has  $\lambda_1 = 4$  and  $\lambda_2 = 5$  then  $\det(A - \lambda I) = (\lambda - 4)(\lambda - 5) = \lambda^2 - 9\lambda + 20$ . Find three matrices that have trace  $a + d = 9$  and determinant 20 and  $\lambda = 4, 5$ .

- 19 A 3 by 3 matrix  $B$  is known to have eigenvalues 0, 1, 2. This information is enough to find three of these (give the answers where possible):
- (a) the rank of  $B$
  - (b) the determinant of  $B^T B$
  - (c) the eigenvalues of  $B^T B$
  - (d) the eigenvalues of  $(B^2 + I)^{-1}$ .
- 20 Choose the last rows of  $A$  and  $C$  to give eigenvalues 4, 7 and 1, 2, 3:

**Companion matrices**

$$A = \begin{bmatrix} 0 & 1 \\ * & * \end{bmatrix} \quad C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ * & * & * \end{bmatrix}.$$

- 21 *The eigenvalues of  $A$  equal the eigenvalues of  $A^T$ .* This is because  $\det(A - \lambda I)$  equals  $\det(A^T - \lambda I)$ . That is true because \_\_\_\_\_. Show by an example that the eigenvectors of  $A$  and  $A^T$  are *not* the same.
- 22 Construct any 3 by 3 Markov matrix  $M$ : positive entries down each column add to 1. Show that  $M^T(1, 1, 1) = (1, 1, 1)$ . By Problem 21,  $\lambda = 1$  is also an eigenvalue of  $M$ . Challenge: A 3 by 3 singular Markov matrix with trace  $\frac{1}{2}$  has what  $\lambda$ 's?
- 23 Find three 2 by 2 matrices that have  $\lambda_1 = \lambda_2 = 0$ . The trace is zero and the determinant is zero.  $A$  might not be the zero matrix but check that  $A^2 = 0$ .
- 24 This matrix is singular with rank one. Find three  $\lambda$ 's and three eigenvectors:

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

- 25 Suppose  $A$  and  $B$  have the same eigenvalues  $\lambda_1, \dots, \lambda_n$  with the same independent eigenvectors  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . Then  $A = B$ . *Reason:* Any vector  $\mathbf{x}$  is a combination  $c_1\mathbf{x}_1 + \dots + c_n\mathbf{x}_n$ . What is  $A\mathbf{x}$ ? What is  $B\mathbf{x}$ ?
- 26 The block  $B$  has eigenvalues 1, 2 and  $C$  has eigenvalues 3, 4 and  $D$  has eigenvalues 5, 7. Find the eigenvalues of the 4 by 4 matrix  $A$ :

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} = \begin{bmatrix} 0 & 1 & 3 & 0 \\ -2 & 3 & 0 & 4 \\ 0 & 0 & 6 & 1 \\ 0 & 0 & 1 & 6 \end{bmatrix}.$$

- 27 Find the rank and the four eigenvalues of  $A$  and  $C$ :

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

- 28 Subtract  $I$  from the previous  $A$ . Find the  $\lambda$ 's and then the determinants of

$$B = A - I = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad C = I - A = \begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 0 & -1 & -1 \\ -1 & -1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}.$$

- 29 (Review) Find the eigenvalues of  $A$ ,  $B$ , and  $C$ :

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}.$$

- 30 When  $a + b = c + d$  show that  $(1, 1)$  is an eigenvector and find both eigenvalues:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

- 31 If we exchange rows 1 and 2 and columns 1 and 2, the eigenvalues don't change. Find eigenvectors of  $A$  and  $B$  for  $\lambda = 11$ . Rank one gives  $\lambda_2 = \lambda_3 = 0$ .

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 6 & 3 \\ 4 & 8 & 4 \end{bmatrix} \quad \text{and} \quad B = PAP^T = \begin{bmatrix} 6 & 3 & 3 \\ 2 & 1 & 1 \\ 8 & 4 & 4 \end{bmatrix}.$$

- 32 Suppose  $A$  has eigenvalues 0, 3, 5 with independent eigenvectors  $\mathbf{u}$ ,  $\mathbf{v}$ ,  $\mathbf{w}$ .

- Give a basis for the nullspace and a basis for the column space.
- Find a particular solution to  $A\mathbf{x} = \mathbf{v} + \mathbf{w}$ . Find all solutions.
- $A\mathbf{x} = \mathbf{u}$  has no solution. If it did then \_\_\_\_\_ would be in the column space.

### Challenge Problems

- 33 Show that  $\mathbf{u}$  is an eigenvector of the rank one  $2 \times 2$  matrix  $A = \mathbf{u}\mathbf{v}^T$ . Find both eigenvalues of  $A$ . Check that  $\lambda_1 + \lambda_2$  agrees with the trace  $u_1v_1 + u_2v_2$ .
- 34 Find the eigenvalues of this permutation matrix  $P$  from  $\det(P - \lambda I) = 0$ . Which vectors are not changed by the permutation? They are eigenvectors for  $\lambda = 1$ . Can you find three more eigenvectors?

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

- 35 There are six 3 by 3 permutation matrices  $P$ . What numbers can be the *determinants* of  $P$ ? What numbers can be *pivots*? What numbers can be the *trace* of  $P$ ? What *four numbers* can be eigenvalues of  $P$ , as in Problem 15?

- 36 (**Heisenberg's Uncertainty Principle**)  $AB - BA = I$  can happen for infinite matrices with  $A = A^T$  and  $B = -B^T$ . Then

$$\mathbf{x}^T \mathbf{x} = \mathbf{x}^T AB\mathbf{x} - \mathbf{x}^T BA\mathbf{x} \leq 2\|A\mathbf{x}\| \|B\mathbf{x}\|.$$

Explain that last step by using the Schwarz inequality  $|\mathbf{u}^T \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$ . Then Heisenberg's inequality says that  $\|A\mathbf{x}\|/\|\mathbf{x}\|$  times  $\|B\mathbf{x}\|/\|\mathbf{x}\|$  is at least  $\frac{1}{2}$ . It is impossible to get the position error and momentum error both very small.

- 37 Find a 2 by 2 rotation matrix (other than  $I$ ) with  $A^3 = I$ . Its eigenvalues must satisfy  $\lambda^3 = 1$ . They can be  $e^{2\pi i/3}$  and  $e^{-2\pi i/3}$ . What are the trace and determinant?
- 38 (a) Find the eigenvalues and eigenvectors of  $A$ . They depend on  $c$ :

$$A = \begin{bmatrix} .4 & 1 - c \\ .6 & c \end{bmatrix}.$$

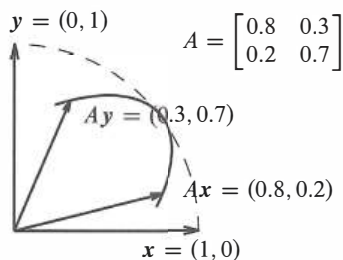
- (b) Show that  $A$  has just one line of eigenvectors when  $c = 1.6$ .

- (c) This is a Markov matrix when  $c = .8$ . Then  $A^n$  will approach what matrix  $A^\infty$ ?

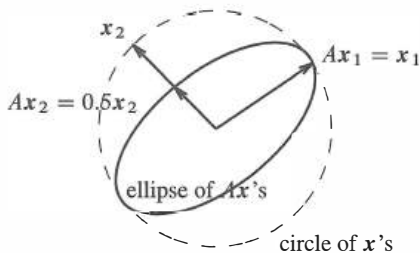
### Eigshow in MATLAB

There is a MATLAB demo (just type **eigshow**), displaying the eigenvalue problem for a 2 by 2 matrix. It starts with the unit vector  $\mathbf{x} = (1, 0)$ . *The mouse makes this vector move around the unit circle.* At the same time the screen shows  $A\mathbf{x}$ , in color and also moving. Possibly  $A\mathbf{x}$  is ahead of  $\mathbf{x}$ . Possibly  $A\mathbf{x}$  is behind  $\mathbf{x}$ . *Sometimes  $A\mathbf{x}$  is parallel to  $\mathbf{x}$ .*

At that parallel moment,  $A\mathbf{x} = \lambda\mathbf{x}$  (at  $\mathbf{x}_1$  and  $\mathbf{x}_2$  in the second figure).



These are not eigenvectors



$A\mathbf{x}$  lines up with  $\mathbf{x}$  at eigenvectors

The eigenvalue  $\lambda$  is the length of  $A\mathbf{x}$ , when the unit eigenvector  $\mathbf{x}$  lines up. The built-in choices for  $A$  illustrate three possibilities: 0, 1, or 2 real vectors where  $A\mathbf{x}$  crosses  $\mathbf{x}$ .

The axes of the ellipse are **singular vectors** in 7.4—and eigenvectors if  $A^T = A$ .

## 6.2 Diagonalizing a Matrix

- 1 The columns of  $AX = X\Lambda$  are  $Ax_k = \lambda_k x_k$ . The eigenvalue matrix  $\Lambda$  is diagonal.
- 2  $n$  independent eigenvectors in  $X$  diagonalize  $A$   $A = X\Lambda X^{-1}$  and  $\Lambda = X^{-1}AX$
- 3 The eigenvector matrix  $X$  also diagonalizes all powers  $A^k$ :  $A^k = X\Lambda^k X^{-1}$
- 4 Solve  $u_{k+1} = Au_k$  by  $u_k = A^k u_0 = X\Lambda^k X^{-1}u_0 = c_1(\lambda_1)^k x_1 + \cdots + c_n(\lambda_n)^k x_n$
- 5 **No equal eigenvalues**  $\Rightarrow X$  is invertible and  $A$  can be diagonalized.  
**Equal eigenvalues**  $\Rightarrow A$  *might* have too few independent eigenvectors. Then  $X^{-1}$  fails.
- 6 Every matrix  $C = B^{-1}AB$  has the **same eigenvalues** as  $A$ . These  $C$ 's are “**similar**” to  $A$ .

When  $x$  is an eigenvector, multiplication by  $A$  is just multiplication by a number  $\lambda$ :  $Ax = \lambda x$ . All the difficulties of matrices are swept away. Instead of an interconnected system, we can follow the eigenvectors separately. It is like having a *diagonal matrix*, with no off-diagonal interconnections. The 100th power of a diagonal matrix is easy.

The point of this section is very direct. *The matrix  $A$  turns into a diagonal matrix  $\Lambda$  when we use the eigenvectors properly.* This is the matrix form of our key idea. We start right off with that one essential computation. The next page explains why  $AX = X\Lambda$ .

**Diagonalization** Suppose the  $n$  by  $n$  matrix  $A$  has  $n$  linearly independent eigenvectors  $x_1, \dots, x_n$ . Put them into the columns of an *eigenvector matrix*  $X$ . Then  $X^{-1}AX$  is the *eigenvalue matrix*  $\Lambda$ :

Eigenvector matrix  $X$   
Eigenvalue matrix  $\Lambda$

$$X^{-1}AX = \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}. \quad (1)$$

The matrix  $A$  is “diagonalized.” We use capital lambda for the eigenvalue matrix, because the small  $\lambda$ 's (the eigenvalues) are on its diagonal.

**Example 1** This  $A$  is triangular so its eigenvalues are on the diagonal:  $\lambda = 1$  and  $\lambda = 6$ .

$$\begin{array}{c} \text{Eigenvectors} \\ \text{go into } X \end{array} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}$$

$$X^{-1} \quad A \quad X = \Lambda$$

In other words  $A = X\Lambda X^{-1}$ . Then watch  $A^2 = X\Lambda X^{-1}X\Lambda X^{-1}$ . So  $A^2$  is  $X\Lambda^2 X^{-1}$ .

*$A^2$  has the same eigenvectors in  $X$  and squared eigenvalues in  $\Lambda^2$ .*

**Why is  $AX = X\Lambda$ ?**  $A$  multiplies its eigenvectors, which are the columns of  $X$ . The first column of  $AX$  is  $Ax_1$ . That is  $\lambda_1 x_1$ . Each column of  $X$  is multiplied by its eigenvalue:

$$\text{A times X} \quad AX = A \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 & \cdots & \lambda_n x_n \end{bmatrix}.$$

The trick is to split this matrix  $AX$  into  $X$  times  $\Lambda$ :

$$\text{X times } \Lambda \quad \begin{bmatrix} \lambda_1 x_1 & \cdots & \lambda_n x_n \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = X\Lambda.$$

Keep those matrices in the right order! Then  $\lambda_1$  multiplies the first column  $x_1$ , as shown. The diagonalization is complete, and we can write  $AX = X\Lambda$  in two good ways:

$$AX = X\Lambda \quad \text{is} \quad X^{-1}AX = \Lambda \quad \text{or} \quad A = X\Lambda X^{-1}. \quad (2)$$

The matrix  $X$  has an inverse, because its columns (the eigenvectors of  $A$ ) were assumed to be linearly independent. *Without  $n$  independent eigenvectors, we can't diagonalize.*

$A$  and  $\Lambda$  have the same eigenvalues  $\lambda_1, \dots, \lambda_n$ . The eigenvectors are different. The job of the original eigenvectors  $x_1, \dots, x_n$  was to diagonalize  $A$ . Those eigenvectors in  $X$  produce  $A = X\Lambda X^{-1}$ . You will soon see their simplicity and importance and meaning. The  $k$ th power will be  $A^k = X\Lambda^k X^{-1}$  which is easy to compute:

$$A^k = (X\Lambda X^{-1})(X\Lambda X^{-1}) \cdots (X\Lambda X^{-1}) = X\Lambda^k X^{-1}.$$

$$\begin{array}{l} \text{Powers of } A \\ \text{Example 1} \end{array} \quad \begin{bmatrix} 1 & 5 \\ 0 & 6 \end{bmatrix}^k = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & 6^k \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 6^k - 1 \\ 0 & 6^k \end{bmatrix} = A^k.$$

With  $k = 1$  we get  $A$ . With  $k = 0$  we get  $A^0 = I$  (and  $\lambda^0 = 1$ ). With  $k = -1$  we get  $A^{-1}$ . You can see how  $A^2 = \begin{bmatrix} 1 & 35 \\ 0 & 36 \end{bmatrix}$  fits that formula when  $k = 2$ .

Here are four small remarks before we use  $\Lambda$  again in Example 2.

**Remark 1** Suppose the eigenvalues  $\lambda_1, \dots, \lambda_n$  are all different. Then it is automatic that the eigenvectors  $x_1, \dots, x_n$  are independent. The eigenvector matrix  $X$  will be *invertible*. **Any matrix that has no repeated eigenvalues can be diagonalized.**

**Remark 2** We can multiply eigenvectors by any nonzero constants.  $A(cx) = \lambda(cx)$  is still true. In Example 1, we can divide  $x = (1, 1)$  by  $\sqrt{2}$  to produce a unit vector.

MATLAB and virtually all other codes produce eigenvectors of length  $\|x\| = 1$ .

**Remark 3** The eigenvectors in  $X$  come in the same order as the eigenvalues in  $\Lambda$ . To reverse the order in  $\Lambda$ , put the eigenvector  $(1, 1)$  before  $(1, 0)$  in  $X$ :

$$\text{New order } 6, 1 \quad \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix} = \Lambda_{\text{new}}$$

To diagonalize  $A$  we *must* use an eigenvector matrix. From  $X^{-1}AX = \Lambda$  we know that  $AX = X\Lambda$ . Suppose the first column of  $X$  is  $x$ . Then the first columns of  $AX$  and  $X\Lambda$  are  $Ax$  and  $\lambda_1 x$ . For those to be equal,  $x$  must be an eigenvector.

**Remark 4** (repeated warning for repeated eigenvalues) Some matrices have too few eigenvectors. *Those matrices cannot be diagonalized.* Here are two examples:

$$\text{Not diagonalizable} \quad A = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Their eigenvalues happen to be 0 and 0. Nothing is special about  $\lambda = 0$ , the problem is the repetition of  $\lambda$ . All eigenvectors of the first matrix are multiples of  $(1, 1)$ :

$$\text{Only one line of eigenvectors} \quad Ax = 0x \quad \text{means} \quad \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{and} \quad x = c \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

There is no second eigenvector, so this unusual matrix  $A$  cannot be diagonalized.

Those matrices are the best examples to test any statement about eigenvectors. In many true-false questions, non-diagonalizable matrices lead to *false*.

Remember that there is no connection between invertibility and diagonalizability:

- **Invertibility** is concerned with the *eigenvalues* ( $\lambda = 0$  or  $\lambda \neq 0$ ).
- **Diagonalizability** is concerned with the *eigenvectors* (too few or enough for  $X$ ).

Each eigenvalue has at least one eigenvector!  $A - \lambda I$  is singular. If  $(A - \lambda I)x = 0$  leads you to  $x = 0$ ,  $\lambda$  is *not* an eigenvalue. Look for a mistake in solving  $\det(A - \lambda I) = 0$ .

**Eigenvectors for  $n$  different  $\lambda$ 's are independent. Then we can diagonalize  $A$ .**

**Independent  $x$  from different  $\lambda$**  Eigenvectors  $x_1, \dots, x_j$  that correspond to distinct (all different) eigenvalues are linearly independent. An  $n$  by  $n$  matrix that has  $n$  different eigenvalues (no repeated  $\lambda$ 's) must be diagonalizable.

**Proof** Suppose  $c_1 x_1 + c_2 x_2 = 0$ . Multiply by  $A$  to find  $c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 = 0$ . Multiply by  $\lambda_2$  to find  $c_1 \lambda_2 x_1 + c_2 \lambda_2 x_2 = 0$ . Now subtract one from the other:

$$\text{Subtraction leaves} \quad (\lambda_1 - \lambda_2)c_1 x_1 = 0. \quad \text{Therefore } c_1 = 0.$$

Since the  $\lambda$ 's are different and  $x_1 \neq 0$ , we are forced to the conclusion that  $c_1 = 0$ . Similarly  $c_2 = 0$ . Only the combination with  $c_1 = c_2 = 0$  gives  $c_1 x_1 + c_2 x_2 = 0$ . So the eigenvectors  $x_1$  and  $x_2$  must be independent.

This proof extends directly to  $j$  eigenvectors. Suppose that  $c_1\mathbf{x}_1 + \cdots + c_j\mathbf{x}_j = \mathbf{0}$ . Multiply by  $A$ , multiply by  $\lambda_j$ , and subtract. This multiplies  $\mathbf{x}_j$  by  $\lambda_j - \lambda_j = 0$ , and  $\mathbf{x}_j$  is gone. Now multiply by  $A$  and by  $\lambda_{j-1}$  and subtract. This removes  $\mathbf{x}_{j-1}$ . Eventually only  $\mathbf{x}_1$  is left:

$$\text{We reach } (\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_j) c_1 \mathbf{x}_1 = \mathbf{0} \quad \text{which forces } c_1 = 0. \quad (3)$$

Similarly every  $c_i = 0$ . When the  $\lambda$ 's are all different, the eigenvectors are independent. A full set of eigenvectors can go into the columns of the eigenvector matrix  $X$ .

**Example 2 Powers of  $A$**  The Markov matrix  $A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}$  in the last section had  $\lambda_1 = 1$  and  $\lambda_2 = .5$ . Here is  $A = X\Lambda X^{-1}$  with those eigenvalues in the diagonal  $\Lambda$ :

$$\text{Markov example} \quad \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix} = X\Lambda X^{-1}.$$

The eigenvectors  $(.6, .4)$  and  $(1, -1)$  are in the columns of  $X$ . They are also the eigenvectors of  $A^2$ . Watch how  $A^2$  has the same  $X$ , and **the eigenvalue matrix of  $A^2$  is  $\Lambda^2$** :

$$\text{Same } X \text{ for } A^2 \quad A^2 = X\Lambda X^{-1}X\Lambda X^{-1} = X\Lambda^2 X^{-1}. \quad (4)$$

Just keep going, and you see why the high powers  $A^k$  approach a “steady state”:

$$\text{Powers of } A \quad A^k = X\Lambda^k X^{-1} = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1^k & 0 \\ 0 & (.5)^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix}.$$

As  $k$  gets larger,  $(.5)^k$  gets smaller. In the limit it disappears completely. That limit is  $A^\infty$ :

$$\text{Limit } k \rightarrow \infty \quad A^\infty = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix} = \begin{bmatrix} .6 & .6 \\ .4 & .4 \end{bmatrix}.$$

The limit has the eigenvector  $\mathbf{x}_1$  in both columns. We saw this  $A^\infty$  on the very first page of Chapter 6. Now we see it coming from powers like  $A^{100} = X\Lambda^{100}X^{-1}$ .

**Question**

When does  $A^k \rightarrow \text{zero matrix}$ ?

**Answer**

All  $|\lambda| < 1$ .

### Similar Matrices: Same Eigenvalues

Suppose the eigenvalue matrix  $\Lambda$  is fixed. As we change the eigenvector matrix  $X$ , we get a whole family of different matrices  $A = X\Lambda X^{-1}$ —all with the same eigenvalues in  $\Lambda$ . All those matrices  $A$  (with the same  $\Lambda$ ) are called **similar**.

This idea extends to matrices that can't be diagonalized. Again we choose one constant matrix  $C$  (not necessarily  $\Lambda$ ). And we look at the whole family of matrices  $A = BCB^{-1}$ , allowing all invertible matrices  $B$ . Again those matrices  $A$  and  $C$  are called **similar**.

We are using  $C$  instead of  $\Lambda$  because  $C$  might not be diagonal. We are using  $B$  instead of  $X$  because the columns of  $B$  might not be eigenvectors. We only require that  $B$  is invertible—its columns can contain any basis for  $\mathbf{R}^n$ . The key fact about similar matrices stays true. **Similar matrices  $A$  and  $C$  have the same eigenvalues.**

**All the matrices  $A = BCB^{-1}$  are “similar.” They all share the eigenvalues of  $C$ .**

**Proof** Suppose  $Cx = \lambda x$ . Then  $BCB^{-1}$  has the same eigenvalue  $\lambda$  with the new eigenvector  $Bx$ :

$$\text{Same } \lambda \quad (BCB^{-1})(Bx) = BCx = B\lambda x = \lambda(Bx). \quad (5)$$

A fixed matrix  $C$  produces a family of similar matrices  $BCB^{-1}$ , allowing all  $B$ . When  $C$  is the identity matrix, the “family” is very small. The only member is  $BIB^{-1} = I$ . The identity matrix is the only diagonalizable matrix with all eigenvalues  $\lambda = 1$ .

The family is larger when  $\lambda = 1$  and 1 *with only one eigenvector* (not diagonalizable). The simplest  $C$  is the *Jordan form*—to be developed in Section 8.3. All the similar  $A$ 's have two parameters  $r$  and  $s$ , not both zero: always determinant = 1 and trace = 2.

$$C = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \text{Jordan form gives } A = BCB^{-1} = \begin{bmatrix} 1-rs & r^2 \\ -s^2 & 1+rs \end{bmatrix}. \quad (6)$$

For an important example I will take eigenvalues  $\lambda = 1$  and 0 (not repeated!). Now the whole family is diagonalizable with the same eigenvalue matrix  $\Lambda$ . We get every 2 by 2 matrix that has eigenvalues 1 and 0. The trace is 1 and the determinant is zero:

$$\text{All similar } \Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{or } A = \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix} \quad \text{or any } A = \frac{xy^T}{x^T y}.$$

The family contains all matrices with  $A^2 = A$ , including  $A = \Lambda$  when  $B = I$ . When  $A$  is symmetric these are also projection matrices. Eigenvalues 1 and 0 make life easy.

## Fibonacci Numbers

We present a famous example, where eigenvalues tell how fast the Fibonacci numbers grow. **Every new Fibonacci number is the sum of the two previous  $F$ 's:**

$$\text{The sequence } 0, 1, 1, 2, 3, 5, 8, 13, \dots \quad \text{comes from} \quad F_{k+2} = F_{k+1} + F_k.$$

These numbers turn up in a fantastic variety of applications. Plants and trees grow in a spiral pattern, and a pear tree has 8 growths for every 3 turns. For a willow those numbers can be 13 and 5. The champion is a sunflower of Daniel O'Connell, which had 233 seeds in 144 loops. Those are the Fibonacci numbers  $F_{13}$  and  $F_{12}$ . Our problem is more basic.

**Problem: Find the Fibonacci number  $F_{100}$**  The slow way is to apply the rule  $F_{k+2} = F_{k+1} + F_k$  one step at a time. By adding  $F_6 = 8$  to  $F_7 = 13$  we reach  $F_8 = 21$ . Eventually we come to  $F_{100}$ . Linear algebra gives a better way.

The key is to begin with a matrix equation  $\mathbf{u}_{k+1} = A\mathbf{u}_k$ . That is a *one-step* rule for vectors, while Fibonacci gave a two-step rule for scalars. We match those rules by putting two Fibonacci numbers into a vector. Then you will see the matrix  $A$ .

$$\text{Let } \mathbf{u}_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}. \text{ The rule } \begin{matrix} F_{k+2} = F_{k+1} + F_k \\ F_{k+1} = F_{k+1} \end{matrix} \text{ is } \mathbf{u}_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{u}_k. \quad (7)$$

**Every step multiplies by  $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ .** After 100 steps we reach  $\mathbf{u}_{100} = A^{100}\mathbf{u}_0$ :

$$\mathbf{u}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \quad \dots, \quad \mathbf{u}_{100} = \begin{bmatrix} F_{101} \\ F_{100} \end{bmatrix}.$$

This problem is just right for eigenvalues. Subtract  $\lambda$  from the diagonal of  $A$ :

$$A - \lambda I = \begin{bmatrix} 1 - \lambda & 1 \\ 1 & -\lambda \end{bmatrix} \quad \text{leads to} \quad \det(A - \lambda I) = \lambda^2 - \lambda - 1.$$

The equation  $\lambda^2 - \lambda - 1 = 0$  is solved by the quadratic formula  $(-b \pm \sqrt{b^2 - 4ac})/2a$ :

$$\text{Eigenvalues} \quad \lambda_1 = \frac{1 + \sqrt{5}}{2} \approx 1.618 \quad \text{and} \quad \lambda_2 = \frac{1 - \sqrt{5}}{2} \approx -.618.$$

These eigenvalues lead to eigenvectors  $\mathbf{x}_1 = (\lambda_1, 1)$  and  $\mathbf{x}_2 = (\lambda_2, 1)$ . Step 2 finds the combination of those eigenvectors that gives  $\mathbf{u}_0 = (1, 0)$ :

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\lambda_1 - \lambda_2} \left( \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} - \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix} \right) \quad \text{or} \quad \mathbf{u}_0 = \frac{\mathbf{x}_1 - \mathbf{x}_2}{\lambda_1 - \lambda_2}. \quad (8)$$

Step 3 multiplies  $\mathbf{u}_0$  by  $A^{100}$  to find  $\mathbf{u}_{100}$ . The eigenvectors  $\mathbf{x}_1$  and  $\mathbf{x}_2$  stay separate! They are multiplied by  $(\lambda_1)^{100}$  and  $(\lambda_2)^{100}$ :

$$\text{100 steps from } \mathbf{u}_0 \quad \mathbf{u}_{100} = \frac{(\lambda_1)^{100}\mathbf{x}_1 - (\lambda_2)^{100}\mathbf{x}_2}{\lambda_1 - \lambda_2}. \quad (9)$$

We want  $F_{100}$  = second component of  $\mathbf{u}_{100}$ . The second components of  $\mathbf{x}_1$  and  $\mathbf{x}_2$  are 1. The difference between  $\lambda_1 = (1 + \sqrt{5})/2$  and  $\lambda_2 = (1 - \sqrt{5})/2$  is  $\sqrt{5}$ . And  $\lambda_2^{100} \approx 0$ .

$$\text{100th Fibonacci number} = \frac{\lambda_1^{100} - \lambda_2^{100}}{\lambda_1 - \lambda_2} = \text{nearest integer to } \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^{100}. \quad (10)$$

Every  $F_k$  is a whole number. The ratio  $F_{101}/F_{100}$  must be very close to the limiting ratio  $(1 + \sqrt{5})/2$ . The Greeks called this number the “golden mean”. For some reason a rectangle with sides 1.618 and 1 looks especially graceful.

Matrix Powers  $A^k$ 

Fibonacci's example is a typical difference equation  $\mathbf{u}_{k+1} = A\mathbf{u}_k$ . *Each step multiplies by  $A$ .* The solution is  $\mathbf{u}_k = A^k\mathbf{u}_0$ . We want to make clear how diagonalizing the matrix gives a quick way to compute  $A^k$  and find  $\mathbf{u}_k$  in three steps.

The eigenvector matrix  $X$  produces  $A = X\Lambda X^{-1}$ . This is a factorization of the matrix, like  $A = LU$  or  $A = QR$ . The new factorization is perfectly suited to computing powers, because *every time  $X^{-1}$  multiplies  $X$  we get  $I$ :*

**Powers of  $A$**   $A^k\mathbf{u}_0 = (X\Lambda X^{-1}) \cdots (X\Lambda X^{-1})\mathbf{u}_0 = X\Lambda^k X^{-1}\mathbf{u}_0$

I will split  $X\Lambda^k X^{-1}\mathbf{u}_0$  into three steps that show how eigenvalues work:

1. Write  $\mathbf{u}_0$  as a combination  $c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n$  of the eigenvectors. Then  $\mathbf{c} = X^{-1}\mathbf{u}_0$ .
2. Multiply each eigenvector  $\mathbf{x}_i$  by  $(\lambda_i)^k$ . Now we have  $\Lambda^k X^{-1}\mathbf{u}_0$ .
3. Add up the pieces  $c_i(\lambda_i)^k\mathbf{x}_i$  to find the solution  $\mathbf{u}_k = A^k\mathbf{u}_0$ . This is  $X\Lambda^k X^{-1}\mathbf{u}_0$ .

**Solution for  $\mathbf{u}_{k+1} = A\mathbf{u}_k$**   $\mathbf{u}_k = A^k\mathbf{u}_0 = c_1(\lambda_1)^k\mathbf{x}_1 + \cdots + c_n(\lambda_n)^k\mathbf{x}_n. \quad (11)$

In matrix language  $A^k$  equals  $(X\Lambda X^{-1})^k$  which is  $X$  times  $\Lambda^k$  times  $X^{-1}$ . In Step 1, the eigenvectors in  $X$  lead to the  $c$ 's in the combination  $\mathbf{u}_0 = c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n$ :

**Step 1**  $\mathbf{u}_0 = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$  This says that  $\mathbf{u}_0 = X\mathbf{c}.$  (12)

The coefficients in Step 1 are  $\mathbf{c} = X^{-1}\mathbf{u}_0$ . Then Step 2 multiplies by  $\Lambda^k$ . The final result  $\mathbf{u}_k = \sum c_i(\lambda_i)^k\mathbf{x}_i$  in Step 3 is the product of  $X$  and  $\Lambda^k$  and  $X^{-1}\mathbf{u}_0$ :

$$A^k\mathbf{u}_0 = X\Lambda^k X^{-1}\mathbf{u}_0 = X\Lambda^k\mathbf{c} = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} (\lambda_1)^k & & \\ & \ddots & \\ & & (\lambda_n)^k \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}. \quad (13)$$

This result is exactly  $\mathbf{u}_k = c_1(\lambda_1)^k\mathbf{x}_1 + \cdots + c_n(\lambda_n)^k\mathbf{x}_n$ . It solves  $\mathbf{u}_{k+1} = A\mathbf{u}_k$ .

**Example 3** Start from  $\mathbf{u}_0 = (1, 0)$ . Compute  $A^k\mathbf{u}_0$  for this faster Fibonacci:

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \quad \text{has} \quad \lambda_1 = 2 \quad \text{and} \quad \mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \lambda_2 = -1 \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

This matrix is like Fibonacci except the rule is changed to  $F_{k+2} = F_{k+1} + 2F_k$ . The new numbers start with 0, 1, 1, 3. They grow faster because of  $\lambda = 2$ .

**Find  $u_k = A^k u_0$  in 3 steps**  $u_0 = c_1 x_1 + c_2 x_2$  and  $u_k = c_1 (\lambda_1)^k x_1 + c_2 (\lambda_2)^k x_2$

**Step 1**  $u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$  so  $c_1 = c_2 = \frac{1}{3}$

**Step 2** Multiply the two parts by  $(\lambda_1)^k = 2^k$  and  $(\lambda_2)^k = (-1)^k$

**Step 3** Combine eigenvectors  $c_1 (\lambda_1)^k x_1$  and  $c_2 (\lambda_2)^k x_2$  into  $u_k$ :

$$u_k = A^k u_0 \quad u_k = \frac{1}{3} 2^k \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{1}{3} (-1)^k \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}.$$

The new number is  $F_k = (2^k - (-1)^k)/3$ . After 0, 1, 1, 3 comes  $F_4 = 15/3 = 5$ .

Behind these numerical examples lies a fundamental idea: **Follow the eigenvectors**. In Section 6.3 this is the crucial link from linear algebra to differential equations ( $\lambda^k$  will become  $e^{\lambda t}$ ). Chapter 8 sees the same idea as “transforming to an eigenvector basis.” The best example of all is a **Fourier series**, built from the eigenvectors  $e^{ikx}$  of  $d/dx$ .

### Nondiagonalizable Matrices (Optional)

Suppose  $\lambda$  is an eigenvalue of  $A$ . We discover that fact in two ways:

1. **Eigenvectors** (geometric) There are nonzero solutions to  $Ax = \lambda x$ .
2. **Eigenvalues** (algebraic) The determinant of  $A - \lambda I$  is zero.

The number  $\lambda$  may be a simple eigenvalue or a multiple eigenvalue, and we want to know its **multiplicity**. Most eigenvalues have multiplicity  $M = 1$  (simple eigenvalues). Then there is a single line of eigenvectors, and  $\det(A - \lambda I)$  does not have a double factor.

For exceptional matrices, an eigenvalue can be **repeated**. Then there are two different ways to count its multiplicity. Always  $GM \leq AM$  for each  $\lambda$ :

1. **(Geometric Multiplicity = GM)** Count the **independent eigenvectors** for  $\lambda$ . Then GM is the dimension of the nullspace of  $A - \lambda I$ .
2. **(Algebraic Multiplicity = AM)** AM counts the **repetitions of  $\lambda$**  among the eigenvalues. Look at the  $n$  roots of  $\det(A - \lambda I) = 0$ .

If  $A$  has  $\lambda = 4, 4, 4$ , then that eigenvalue has  $AM = 3$  and  $GM = 1, 2$ , or  $3$ .

The following matrix  $A$  is the standard example of trouble. Its eigenvalue  $\lambda = 0$  is repeated. It is a double eigenvalue ( $AM = 2$ ) with only one eigenvector ( $GM = 1$ ).

$$\begin{array}{ll} AM = 2 & A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ has } \det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ 0 & -\lambda \end{vmatrix} = \lambda^2. & \lambda = 0, 0 \text{ but} \\ GM = 1 & & \text{1 eigenvector} \end{array}$$

There “should” be two eigenvectors, because  $\lambda^2 = 0$  has a double root. The double factor  $\lambda^2$  makes  $\text{AM} = 2$ . But there is only one eigenvector  $\mathbf{x} = (1, 0)$  and  $\text{GM} = 1$ . ***This shortage of eigenvectors when GM is below AM means that  $A$  is not diagonalizable.***

These three matrices all have the same shortage of eigenvectors. Their repeated eigenvalue is  $\lambda = 5$ . Traces are 10 and determinants are 25:

$$A = \begin{bmatrix} 5 & 1 \\ 0 & 5 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 6 & -1 \\ 1 & 4 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 7 & 2 \\ -2 & 3 \end{bmatrix}.$$

Those all have  $\det(A - \lambda I) = (\lambda - 5)^2$ . The algebraic multiplicity is  $\text{AM} = 2$ . But each  $A - 5I$  has rank  $r = 1$ . The geometric multiplicity is  $\text{GM} = 1$ . There is only one line of eigenvectors for  $\lambda = 5$ , and these matrices are not diagonalizable.

## ■ REVIEW OF THE KEY IDEAS ■

1. If  $A$  has  $n$  independent eigenvectors  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , they go into the columns of  $X$ .

$$A \text{ is diagonalized by } X \quad X^{-1}AX = \Lambda \quad \text{and} \quad A = X\Lambda X^{-1}.$$

2. The powers of  $A$  are  $A^k = X\Lambda^k X^{-1}$ . The eigenvectors in  $X$  are unchanged.
3. The eigenvalues of  $A^k$  are  $(\lambda_1)^k, \dots, (\lambda_n)^k$  in the matrix  $\Lambda^k$ .
4. The solution to  $\mathbf{u}_{k+1} = A\mathbf{u}_k$  starting from  $\mathbf{u}_0$  is  $\mathbf{u}_k = A^k\mathbf{u}_0 = X\Lambda^k X^{-1}\mathbf{u}_0$ :

$$\mathbf{u}_k = c_1(\lambda_1)^k \mathbf{x}_1 + \dots + c_n(\lambda_n)^k \mathbf{x}_n \quad \text{provided} \quad \mathbf{u}_0 = c_1 \mathbf{x}_1 + \dots + c_n \mathbf{x}_n.$$

That shows Steps 1, 2, 3 ( $c$ 's from  $X^{-1}\mathbf{u}_0$ ,  $\lambda^k$  from  $\Lambda^k$ , and  $\mathbf{x}$ 's from  $X$ )

5.  $A$  is diagonalizable if every eigenvalue has enough eigenvectors ( $\text{GM} = \text{AM}$ ).

## ■ WORKED EXAMPLES ■

**6.2 A** The **Lucas numbers** are like the Fibonacci numbers except they start with  $L_1 = 1$  and  $L_2 = 3$ . Using the same rule  $L_{k+2} = L_{k+1} + L_k$ , the next Lucas numbers are 4, 7, 11, 18. Show that the Lucas number  $L_{100}$  is  $\lambda_1^{100} + \lambda_2^{100}$ .

**Solution**  $\mathbf{u}_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{u}_k$  is the same as for Fibonacci, because  $L_{k+2} = L_{k+1} + L_k$  is the same rule (with different starting values). The equation becomes a 2 by 2 system:

Let  $\mathbf{u}_k = \begin{bmatrix} L_{k+1} \\ L_k \end{bmatrix}$ . The rule  $\begin{matrix} L_{k+2} = L_{k+1} + L_k \\ L_{k+1} = L_{k+1} \end{matrix}$  is  $\mathbf{u}_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \mathbf{u}_k$ .

The eigenvalues and eigenvectors of  $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$  still come from  $\lambda^2 = \lambda + 1$ :

$$\lambda_1 = \frac{1 + \sqrt{5}}{2} \quad \text{and} \quad \mathbf{x}_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} \quad \lambda_2 = \frac{1 - \sqrt{5}}{2} \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}.$$

Now solve  $c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 = \mathbf{u}_1 = (3, 1)$ . The solution is  $c_1 = \lambda_1$  and  $c_2 = \lambda_2$ . Check:

$$\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 = \begin{bmatrix} \lambda_1^2 + \lambda_2^2 \\ \lambda_1 + \lambda_2 \end{bmatrix} = \begin{bmatrix} \text{trace of } A^2 \\ \text{trace of } A \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \mathbf{u}_1$$

$\mathbf{u}_{100} = A^{99} \mathbf{u}_1$  tells us the Lucas numbers  $(L_{101}, L_{100})$ . The second components of the eigenvectors  $\mathbf{x}_1$  and  $\mathbf{x}_2$  are 1, so the second component of  $\mathbf{u}_{100}$  is the answer we want:

$$\text{Lucas number} \quad L_{100} = c_1 \lambda_1^{99} + c_2 \lambda_2^{99} = \lambda_1^{100} + \lambda_2^{100}.$$

Lucas starts faster than Fibonacci, and ends up larger by a factor near  $\sqrt{5}$ .

**6.2 B** Find the inverse and the eigenvalues and the determinant of this matrix  $A$ :

$$A = 5 * \mathbf{eye}(4) - \mathbf{ones}(4) = \begin{bmatrix} 4 & -1 & -1 & -1 \\ -1 & 4 & -1 & -1 \\ -1 & -1 & 4 & -1 \\ -1 & -1 & -1 & 4 \end{bmatrix}.$$

Describe an eigenvector matrix  $X$  that gives  $X^{-1}AX = \Lambda$ .

**Solution** What are the eigenvalues of the all-ones matrix? Its rank is certainly 1, so three eigenvalues are  $\lambda = 0, 0, 0$ . Its trace is 4, so the other eigenvalue is  $\lambda = 4$ . Subtract this all-ones matrix from  $5I$  to get our matrix  $A$ :

**Subtract the eigenvalues 4, 0, 0, 0 from 5, 5, 5, 5. The eigenvalues of  $A$  are 1, 5, 5, 5.**

The determinant of  $A$  is 125, the product of those four eigenvalues. The eigenvector for  $\lambda = 1$  is  $\mathbf{x} = (1, 1, 1, 1)$  or  $(c, c, c, c)$ . The other eigenvectors are perpendicular to  $\mathbf{x}$  (since  $A$  is symmetric). The nicest eigenvector matrix  $X$  is the symmetric orthogonal **Hadamard matrix  $H$** . The factor  $\frac{1}{2}$  produces unit column vectors.

$$\text{Orthonormal eigenvectors} \quad X = H = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} = H^T = H^{-1}.$$

The eigenvalues of  $A^{-1}$  are  $1, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}$ . The eigenvectors are not changed so  $A^{-1} = H\Lambda^{-1}H^{-1}$ . The inverse matrix is surprisingly neat:

$$A^{-1} = \frac{1}{5} * (\text{eye}(4) + \text{ones}(4)) = \frac{1}{5} \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{bmatrix}$$

$A$  is a rank-one change from  $5I$ . So  $A^{-1}$  is a rank-one change from  $I/5$ .

In a graph with 5 nodes, the determinant 125 counts the “spanning trees” (trees that touch all nodes). *Trees have no loops* (graphs and trees are in Section 10.1).

With 6 nodes, the matrix  $6 * \text{eye}(5) - \text{ones}(5)$  has the five eigenvalues 1, 6, 6, 6, 6.

## Problem Set 6.2

**Questions 1–7 are about the eigenvalue and eigenvector matrices  $\Lambda$  and  $X$ .**

- 1 (a) Factor these two matrices into  $A = X\Lambda X^{-1}$ :

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 \\ 3 & 3 \end{bmatrix}.$$

- (b) If  $A = X\Lambda X^{-1}$  then  $A^3 = ( ) ( ) ( )$  and  $A^{-1} = ( ) ( ) ( )$ .
- 2 If  $A$  has  $\lambda_1 = 2$  with eigenvector  $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$  and  $\lambda_2 = 5$  with  $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ , use  $X\Lambda X^{-1}$  to find  $A$ . No other matrix has the same  $\lambda$ 's and  $\mathbf{x}$ 's.
- 3 Suppose  $A = X\Lambda X^{-1}$ . What is the eigenvalue matrix for  $A + 2I$ ? What is the eigenvector matrix? Check that  $A + 2I = ( ) ( ) ( )^{-1}$ .
- 4 True or false: If the columns of  $X$  (eigenvectors of  $A$ ) are linearly independent, then
- (a)  $A$  is invertible      (b)  $A$  is diagonalizable
- (c)  $X$  is invertible      (d)  $X$  is diagonalizable.
- 5 If the eigenvectors of  $A$  are the columns of  $I$ , then  $A$  is a \_\_\_\_\_ matrix. If the eigenvector matrix  $X$  is triangular, then  $X^{-1}$  is triangular. Prove that  $A$  is also triangular.
- 6 Describe all matrices  $X$  that diagonalize this matrix  $A$  (find all eigenvectors):

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

Then describe all matrices that diagonalize  $A^{-1}$ .

- 7 Write down the most general matrix that has eigenvectors  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  and  $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ .

**Questions 8–10 are about Fibonacci and Gibonacci numbers.**

- 8** Diagonalize the Fibonacci matrix by completing  $X^{-1}$ :

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} & \\ & \end{bmatrix}.$$

Do the multiplication  $X\Lambda^k X^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$  to find its second component. This is the  $k$ th Fibonacci number  $F_k = (\lambda_1^k - \lambda_2^k)/(\lambda_1 - \lambda_2)$ .

- 9** Suppose  $G_{k+2}$  is the *average* of the two previous numbers  $G_{k+1}$  and  $G_k$ :

$$\begin{array}{l} G_{k+2} = \frac{1}{2}G_{k+1} + \frac{1}{2}G_k \\ G_{k+1} = G_k \end{array} \quad \text{is} \quad \begin{bmatrix} G_{k+2} \\ G_{k+1} \end{bmatrix} = \begin{bmatrix} A \\ \end{bmatrix} \begin{bmatrix} G_{k+1} \\ G_k \end{bmatrix}.$$

- Find the eigenvalues and eigenvectors of  $A$ .
  - Find the limit as  $n \rightarrow \infty$  of the matrices  $A^n = X\Lambda^n X^{-1}$ .
  - If  $G_0 = 0$  and  $G_1 = 1$  show that the Gibonacci numbers approach  $\frac{2}{3}$ .
- 10** Prove that every third Fibonacci number in  $0, 1, 1, 2, 3, \dots$  is even.

**Questions 11–14 are about diagonalizability.**

- 11** True or false: If the eigenvalues of  $A$  are 2, 2, 5 then the matrix is certainly
- invertible
  - diagonalizable
  - not diagonalizable.
- 12** True or false: If the only eigenvectors of  $A$  are multiples of  $(1, 4)$  then  $A$  has
- no inverse
  - a repeated eigenvalue
  - no diagonalization  $X\Lambda X^{-1}$ .
- 13** Complete these matrices so that  $\det A = 25$ . Then check that  $\lambda = 5$  is repeated—the trace is 10 so the determinant of  $A - \lambda I$  is  $(\lambda - 5)^2$ . Find an eigenvector with  $Ax = 5x$ . These matrices will not be diagonalizable because there is no second line of eigenvectors.

$$A = \begin{bmatrix} 8 & \\ & 2 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 9 & 4 \\ & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 10 & 5 \\ -5 & \end{bmatrix}$$

- 14** The matrix  $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$  is not diagonalizable because the rank of  $A - 3I$  is \_\_\_\_\_. Change one entry to make  $A$  diagonalizable. Which entries could you change?

**Questions 15–19 are about powers of matrices.**

- 15**  $A^k = X\Lambda^k X^{-1}$  approaches the zero matrix as  $k \rightarrow \infty$  if and only if every  $\lambda$  has absolute value less than \_\_\_\_\_. Which of these matrices has  $A^k \rightarrow 0$ ?

$$A_1 = \begin{bmatrix} .6 & .9 \\ .4 & .1 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} .6 & .9 \\ .1 & .6 \end{bmatrix}.$$

- 16 (Recommended) Find  $\Lambda$  and  $X$  to diagonalize  $A_1$  in Problem 15. What is the limit of  $\Lambda^k$  as  $k \rightarrow \infty$ ? What is the limit of  $X\Lambda^kX^{-1}$ ? In the columns of this limiting matrix you see the \_\_\_\_\_.

- 17 Find  $\Lambda$  and  $X$  to diagonalize  $A_2$  in Problem 15. What is  $(A_2)^{10}u_0$  for these  $u_0$ ?

$$u_0 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{and} \quad u_0 = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad \text{and} \quad u_0 = \begin{bmatrix} 6 \\ 0 \end{bmatrix}.$$

- 18 Diagonalize  $A$  and compute  $X\Lambda^kX^{-1}$  to prove this formula for  $A^k$ :

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \quad \text{has} \quad A^k = \frac{1}{2} \begin{bmatrix} 1 + 3^k & 1 - 3^k \\ 1 - 3^k & 1 + 3^k \end{bmatrix}.$$

- 19 Diagonalize  $B$  and compute  $X\Lambda^kX^{-1}$  to prove this formula for  $B^k$ :

$$B = \begin{bmatrix} 5 & 1 \\ 0 & 4 \end{bmatrix} \quad \text{has} \quad B^k = \begin{bmatrix} 5^k & 5^k - 4^k \\ 0 & 4^k \end{bmatrix}.$$

- 20 Suppose  $A = X\Lambda X^{-1}$ . Take determinants to prove  $\det A = \det \Lambda = \lambda_1 \lambda_2 \cdots \lambda_n$ . This quick proof only works when  $A$  can be \_\_\_\_\_.

- 21 Show that  $\text{trace } XY = \text{trace } YX$ , by adding the diagonal entries of  $XY$  and  $YX$ :

$$X = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} q & r \\ s & t \end{bmatrix}.$$

Now choose  $Y$  to be  $\Lambda X^{-1}$ . Then  $X\Lambda X^{-1}$  has the same trace as  $\Lambda X^{-1}X = \Lambda$ . This proves that *the trace of  $A$  equals the trace of  $\Lambda = \text{sum of the eigenvalues}$ .*

- 22  $AB - BA = I$  is impossible since the left side has trace = \_\_\_\_\_. But find an elimination matrix so that  $A = E$  and  $B = E^T$  give

$$AB - BA = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{which has trace zero.}$$

- 23 If  $A = X\Lambda X^{-1}$ , diagonalize the block matrix  $B = \begin{bmatrix} A & 0 \\ 0 & 2A \end{bmatrix}$ . Find its eigenvalue and eigenvector (block) matrices.

- 24 Consider all 4 by 4 matrices  $A$  that are diagonalized by the same fixed eigenvector matrix  $X$ . Show that the  $A$ 's form a subspace ( $cA$  and  $A_1 + A_2$  have this same  $X$ ). What is this subspace when  $X = I$ ? What is its dimension?

- 25 Suppose  $A^2 = A$ . On the left side  $A$  multiplies each column of  $A$ . Which of our four subspaces contains eigenvectors with  $\lambda = 1$ ? Which subspace contains eigenvectors with  $\lambda = 0$ ? From the dimensions of those subspaces,  $A$  has a full set of independent eigenvectors. So a matrix with  $A^2 = A$  can be diagonalized.

- 26 (Recommended) Suppose  $Ax = \lambda x$ . If  $\lambda = 0$  then  $x$  is in the nullspace. If  $\lambda \neq 0$  then  $x$  is in the column space. Those spaces have dimensions  $(n - r) + r = n$ . So why doesn't every square matrix have  $n$  linearly independent eigenvectors?
- 27 The eigenvalues of  $A$  are 1 and 9, and the eigenvalues of  $B$  are  $-1$  and 9:

$$A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 4 & 5 \\ 5 & 4 \end{bmatrix}.$$

Find a matrix square root of  $A$  from  $R = X\sqrt{\Lambda}X^{-1}$ . Why is there no real matrix square root of  $B$ ?

- 28 If  $A$  and  $B$  have the same  $\lambda$ 's with the same independent eigenvectors, their factorizations into \_\_\_\_\_ are the same. So  $A = B$ .
- 29 Suppose the same  $X$  diagonalizes both  $A$  and  $B$ . They have the *same eigenvectors* in  $A = X\Lambda_1X^{-1}$  and  $B = X\Lambda_2X^{-1}$ . Prove that  $AB = BA$ .
- 30 (a) If  $A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$  then the determinant of  $A - \lambda I$  is  $(\lambda - a)(\lambda - d)$ . Check the "Cayley-Hamilton Theorem" that  $(A - aI)(A - dI) = \text{zero matrix}$ .
- (b) Test the Cayley-Hamilton Theorem on Fibonacci's  $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ . The theorem predicts that  $A^2 - A - I = 0$ , since the polynomial  $\det(A - \lambda I)$  is  $\lambda^2 - \lambda - 1$ .
- 31 Substitute  $A = X\Lambda X^{-1}$  into the product  $(A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_n I)$  and explain why this produces the zero matrix. We are substituting the matrix  $A$  for the number  $\lambda$  in the polynomial  $p(\lambda) = \det(A - \lambda I)$ . The **Cayley-Hamilton Theorem** says that this product is always  $p(A) = \text{zero matrix}$ , even if  $A$  is not diagonalizable.
- 32 If  $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$  and  $AB = BA$ , show that  $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  is also a diagonal matrix.  $B$  has the same eigen \_\_\_\_\_ as  $A$  but different eigen \_\_\_\_\_. These diagonal matrices  $B$  form a two-dimensional subspace of matrix space.  $AB - BA = 0$  gives four equations for the unknowns  $a, b, c, d$ —find the rank of the 4 by 4 matrix.
- 33 The powers  $A^k$  approach zero if all  $|\lambda_i| < 1$  and they blow up if any  $|\lambda_i| > 1$ . Peter Lax gives these striking examples in his book *Linear Algebra*:

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 2 \\ -5 & -3 \end{bmatrix} \quad C = \begin{bmatrix} 5 & 7 \\ -3 & -4 \end{bmatrix} \quad D = \begin{bmatrix} 5 & 6.9 \\ -3 & -4 \end{bmatrix}$$

$$\|A^{1024}\| > 10^{700} \quad B^{1024} = I \quad C^{1024} = -C \quad \|D^{1024}\| < 10^{-78}$$

Find the eigenvalues  $\lambda = e^{i\theta}$  of  $B$  and  $C$  to show  $B^4 = I$  and  $C^3 = -I$ .

### Challenge Problems

- 34 The  $n$ th power of rotation through  $\theta$  is rotation through  $n\theta$ :

$$A^n = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}^n = \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}.$$

Prove that neat formula by diagonalizing  $A = X\Lambda X^{-1}$ . The eigenvectors (columns of  $X$ ) are  $(1, i)$  and  $(i, 1)$ . You need to know Euler's formula  $e^{i\theta} = \cos \theta + i \sin \theta$ .

- 35 The transpose of  $A = X\Lambda X^{-1}$  is  $A^T = (X^{-1})^T \Lambda X^T$ . The eigenvectors in  $A^T \mathbf{y} = \lambda \mathbf{y}$  are the columns of that matrix  $(X^{-1})^T$ . They are often called **left eigenvectors of  $A$** , because  $\mathbf{y}^T A = \lambda \mathbf{y}^T$ . How do you multiply matrices to find this formula for  $A$ ?

**Sum of rank-1 matrices**  $A = X\Lambda X^{-1} = \lambda_1 \mathbf{x}_1 \mathbf{y}_1^T + \cdots + \lambda_n \mathbf{x}_n \mathbf{y}_n^T.$

- 36 The inverse of  $A = \mathbf{eye}(n) + \mathbf{ones}(n)$  is  $A^{-1} = \mathbf{eye}(n) + C * \mathbf{ones}(n)$ . Multiply  $AA^{-1}$  to find that number  $C$  (depending on  $n$ ).
- 37 Suppose  $A_1$  and  $A_2$  are  $n$  by  $n$  invertible matrices. What matrix  $B$  shows that  $A_2 A_1 = B(A_1 A_2)B^{-1}$ ? Then  $A_2 A_1$  is similar to  $A_1 A_2$ : *same eigenvalues*.
- 38 **When is a matrix  $A$  similar to its eigenvalue matrix  $\Lambda$ ?**  
 $A$  and  $\Lambda$  always have the same eigenvalues. But similarity requires a matrix  $B$  with  $A = B\Lambda B^{-1}$ . Then  $B$  is the \_\_\_\_\_ matrix and  $A$  must have  $n$  independent \_\_\_\_\_.
- 39 (Pavel Grinfeld) Without writing down any calculations, can you find the eigenvalues of this matrix? Can you find the 2017th power  $A^{2017}$ ?

$$A = \begin{bmatrix} 110 & 55 & -164 \\ 42 & 21 & -62 \\ 88 & 44 & -131 \end{bmatrix}.$$

**If  $A$  is  $m$  by  $n$  and  $B$  is  $n$  by  $m$ , then  $AB$  and  $BA$  have same nonzero eigenvalues.**

*Proof.* Start with this identity between square matrices (easily checked). The first and third matrices are inverses. The “size matrix” shows the shapes of all blocks.

$$\begin{bmatrix} I & -A \\ 0 & I \end{bmatrix} \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \begin{bmatrix} I & A \\ 0 & I \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix} \quad \begin{bmatrix} m \times m & m \times n \\ n \times m & n \times n \end{bmatrix}$$

This equation  $D^{-1}ED = F$  says  **$F$  is similar to  $E$** —they have the *same*  $m+n$  eigenvalues.

$$E = \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \text{ has the } m \text{ eigenvalues of } AB, \text{ plus } n \text{ zeros}$$

$$F = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix} \text{ has the } n \text{ eigenvalues of } BA, \text{ plus } m \text{ zeros}$$

So  $AB$  and  $BA$  have the **same eigenvalues** except for  $|n - m|$  zeros. Wow.

If  $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$  and  $B = A^T$  then  $A^T A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$  (notice  $\lambda = 2$  and 0) and  $AA^T = \begin{bmatrix} 2 \end{bmatrix}$ .

## 6.3 Systems of Differential Equations

- 1 If  $Ax = \lambda x$  then  $\mathbf{u}(t) = e^{\lambda t} \mathbf{x}$  will solve  $\frac{d\mathbf{u}}{dt} = A\mathbf{u}$ . Each  $\lambda$  and  $\mathbf{x}$  give a solution  $e^{\lambda t} \mathbf{x}$ .
- 2 If  $A = X\Lambda X^{-1}$  then  $\mathbf{u}(t) = e^{At} \mathbf{u}(0) = X e^{\Lambda t} X^{-1} \mathbf{u}(0) = c_1 e^{\lambda_1 t} \mathbf{x}_1 + \cdots + c_n e^{\lambda_n t} \mathbf{x}_n$ .
- 3  $A$  is **stable** and  $\mathbf{u}(t) \rightarrow \mathbf{0}$  and  $e^{At} \rightarrow 0$  when all eigenvalues of  $A$  have real part  $< 0$ .
- 4 **Matrix exponential**  $e^{At} = I + At + \cdots + (At)^n/n! + \cdots = X e^{\Lambda t} X^{-1}$  if  $A$  is diagonalizable.
- 5 **Second order equation**  $u'' + Bu' + Cu = 0$  is equivalent to  $\begin{bmatrix} u \\ u' \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ -C & -B \end{bmatrix} \begin{bmatrix} u \\ u' \end{bmatrix}$ .  
**First order system**

Eigenvalues and eigenvectors and  $A = X\Lambda X^{-1}$  are perfect for matrix powers  $A^k$ . They are also perfect for differential equations  $d\mathbf{u}/dt = A\mathbf{u}$ . This section is mostly linear algebra, but to read it you need one fact from calculus: *The derivative of  $e^{\lambda t}$  is  $\lambda e^{\lambda t}$ .* The whole point of the section is this: **To convert constant-coefficient differential equations into linear algebra.**

The ordinary equations  $\frac{du}{dt} = u$  and  $\frac{du}{dt} = \lambda u$  are solved by exponentials:

$$\frac{du}{dt} = u \text{ produces } u(t) = Ce^t \quad \frac{du}{dt} = \lambda u \text{ produces } u(t) = Ce^{\lambda t} \quad (1)$$

At time  $t = 0$  those solutions include  $e^0 = 1$ . So they both reduce to  $u(0) = C$ . This “initial value” tells us the right choice for  $C$ . **The solutions that start from the number  $u(0)$  at time  $t = 0$  are  $u(t) = u(0)e^t$  and  $u(t) = u(0)e^{\lambda t}$ .**

We just solved a 1 by 1 problem. Linear algebra moves to  $n$  by  $n$ . The unknown is a vector  $\mathbf{u}$  (now boldface). It starts from the initial vector  $\mathbf{u}(0)$ , which is given. The  $n$  equations contain a square matrix  $A$ . We expect  $n$  exponents  $e^{\lambda t}$  in  $\mathbf{u}(t)$ , from  $n$   $\lambda$ 's:

$$\text{System of } n \text{ equations} \quad \frac{d\mathbf{u}}{dt} = A\mathbf{u} \quad \text{starting from the vector } \mathbf{u}(0) = \begin{bmatrix} u_1(0) \\ \vdots \\ u_n(0) \end{bmatrix} \text{ at } t = 0. \quad (2)$$

These differential equations are *linear*. If  $\mathbf{u}(t)$  and  $\mathbf{v}(t)$  are solutions, so is  $C\mathbf{u}(t) + D\mathbf{v}(t)$ . We will need  $n$  constants like  $C$  and  $D$  to match the  $n$  components of  $\mathbf{u}(0)$ . Our first job is to find  $n$  “pure exponential solutions”  $\mathbf{u} = e^{\lambda t} \mathbf{x}$  by using  $A\mathbf{x} = \lambda \mathbf{x}$ .

Notice that  $A$  is a *constant* matrix. In other linear equations,  $A$  changes as  $t$  changes. In nonlinear equations,  $A$  changes as  $\mathbf{u}$  changes. We don't have those difficulties,  $d\mathbf{u}/dt = A\mathbf{u}$  is “linear with constant coefficients”. Those and only those are the differential equations that we will convert directly to linear algebra. Here is the key:

*Solve linear constant coefficient equations by exponentials  $e^{\lambda t} \mathbf{x}$ , when  $A\mathbf{x} = \lambda \mathbf{x}$ .*

Solution of  $du/dt = Au$ 

Our pure exponential solution will be  $e^{\lambda t}$  times a fixed vector  $x$ . You may guess that  $\lambda$  is an eigenvalue of  $A$ , and  $x$  is *the eigenvector*. Substitute  $u(t) = e^{\lambda t}x$  into the equation  $du/dt = Au$  to prove you are right. The factor  $e^{\lambda t}$  will cancel to leave  $\lambda x = Ax$ :

$$\text{Choose } u = e^{\lambda t}x \quad \frac{du}{dt} = \lambda e^{\lambda t}x \quad \text{agrees with} \quad Au = Ae^{\lambda t}x \quad (3)$$

All components of this special solution  $u = e^{\lambda t}x$  share the same  $e^{\lambda t}$ . The solution grows when  $\lambda > 0$ . It decays when  $\lambda < 0$ . If  $\lambda$  is a complex number, its real part decides growth or decay. The imaginary part  $\omega$  gives oscillation  $e^{i\omega t}$  like a sine wave.

**Example 1** Solve  $\frac{du}{dt} = Au = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} u$  starting from  $u(0) = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ .

This is a vector equation for  $u$ . It contains two scalar equations for the components  $y$  and  $z$ . They are “coupled together” because the matrix  $A$  is not diagonal:

$$\frac{du}{dt} = Au \quad \frac{d}{dt} \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} \quad \text{means that} \quad \frac{dy}{dt} = z \quad \text{and} \quad \frac{dz}{dt} = y.$$

The idea of eigenvectors is to combine those equations in a way that gets back to 1 by 1 problems. The combinations  $y + z$  and  $y - z$  will do it. Add and subtract equations:

$$\frac{d}{dt}(y + z) = z + y \quad \text{and} \quad \frac{d}{dt}(y - z) = -(y - z).$$

The combination  $y + z$  grows like  $e^t$ , because it has  $\lambda = 1$ . The combination  $y - z$  decays like  $e^{-t}$ , because it has  $\lambda = -1$ . Here is the point: We don’t have to juggle the original equations  $du/dt = Au$ , looking for these special combinations. The eigenvectors and eigenvalues of  $A$  will do it for us.

This matrix  $A$  has eigenvalues 1 and  $-1$ . The eigenvectors  $x$  are  $(1, 1)$  and  $(1, -1)$ . The pure exponential solutions  $u_1$  and  $u_2$  take the form  $e^{\lambda t}x$  with  $\lambda_1 = 1$  and  $\lambda_2 = -1$ :

$$u_1(t) = e^{\lambda_1 t}x_1 = e^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad u_2(t) = e^{\lambda_2 t}x_2 = e^{-t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}. \quad (4)$$

Notice: These  $u$ ’s satisfy  $Au_1 = u_1$  and  $Au_2 = -u_2$ , just like  $x_1$  and  $x_2$ . The factors  $e^t$  and  $e^{-t}$  change with time. Those factors give  $du_1/dt = u_1 = Au_1$  and  $du_2/dt = -u_2 = Au_2$ . **We have two solutions to  $du/dt = Au$ .** To find all other solutions, **multiply those special solutions by any numbers  $C$  and  $D$  and add:**

**Complete solution** 
$$u(t) = Ce^t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + De^{-t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} Ce^t + De^{-t} \\ Ce^t - De^{-t} \end{bmatrix}. \quad (5)$$

With these two constants  $C$  and  $D$ , we can match any starting vector  $\mathbf{u}(0) = (u_1(0), u_2(0))$ . Set  $t = 0$  and  $e^0 = 1$ . Example 1 asked for the initial value to be  $\mathbf{u}(0) = (4, 2)$ :

$$\mathbf{u}(0) \text{ decides } C, D \quad C \begin{bmatrix} 1 \\ 1 \end{bmatrix} + D \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \quad \text{yields } C = 3 \quad \text{and } D = 1.$$

With  $C = 3$  and  $D = 1$  in the solution (5), the initial value problem is completely solved.

The same three steps that solved  $\mathbf{u}_{k+1} = A\mathbf{u}_k$  now solve  $d\mathbf{u}/dt = A\mathbf{u}$ :

1. Write  $\mathbf{u}(0)$  as a **combination**  $c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n$  **of the eigenvectors of  $A$** .
2. Multiply each eigenvector  $\mathbf{x}_i$  by **its growth factor**  $e^{\lambda_i t}$ .
3. The solution is the same combination of those pure solutions  $e^{\lambda_i t}\mathbf{x}_i$ :

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u}$$

$$\mathbf{u}(t) = c_1 e^{\lambda_1 t} \mathbf{x}_1 + \cdots + c_n e^{\lambda_n t} \mathbf{x}_n. \quad (6)$$

*Not included:* If two  $\lambda$ 's are equal, with only one eigenvector, another solution is needed. (It will be  $te^{\lambda t}\mathbf{x}$ .) Step 1 needs to diagonalize  $A = X\Lambda X^{-1}$ : a basis of  $n$  eigenvectors.

**Example 2** Solve  $d\mathbf{u}/dt = A\mathbf{u}$  knowing the eigenvalues  $\lambda = 1, 2, 3$  of  $A$ :

**Typical example**

**Equation for  $\mathbf{u}$**

**Initial condition  $\mathbf{u}(0)$**

$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{u} \quad \text{starting from } \mathbf{u}(0) = \begin{bmatrix} 9 \\ 7 \\ 4 \end{bmatrix}.$$

The eigenvectors are  $\mathbf{x}_1 = (1, 0, 0)$  and  $\mathbf{x}_2 = (1, 1, 0)$  and  $\mathbf{x}_3 = (1, 1, 1)$ .

**Step 1** The vector  $\mathbf{u}(0) = (9, 7, 4)$  is  $2\mathbf{x}_1 + 3\mathbf{x}_2 + 4\mathbf{x}_3$ . Thus  $(c_1, c_2, c_3) = (2, 3, 4)$ .

**Step 2** The factors  $e^{\lambda t}$  give exponential solutions  $e^t\mathbf{x}_1$  and  $e^{2t}\mathbf{x}_2$  and  $e^{3t}\mathbf{x}_3$ .

**Step 3** The combination that starts from  $\mathbf{u}(0)$  is  $\mathbf{u}(t) = 2e^t\mathbf{x}_1 + 3e^{2t}\mathbf{x}_2 + 4e^{3t}\mathbf{x}_3$ .

The coefficients 2, 3, 4 came from solving the linear equation  $c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + c_3\mathbf{x}_3 = \mathbf{u}(0)$ :

$$\begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ 4 \end{bmatrix} \quad \text{which is } X\mathbf{c} = \mathbf{u}(0). \quad (7)$$

You now have the basic idea—how to solve  $d\mathbf{u}/dt = A\mathbf{u}$ . The rest of this section goes further. We solve equations that contain *second* derivatives, because they arise so often in applications. We also decide whether  $\mathbf{u}(t)$  approaches zero or blows up or just oscillates.

At the end comes the **matrix exponential**  $e^{At}$ . The short formula  $e^{At}\mathbf{u}(0)$  solves the equation  $d\mathbf{u}/dt = A\mathbf{u}$  in the same way that  $A^k\mathbf{u}_0$  solves the equation  $\mathbf{u}_{k+1} = A\mathbf{u}_k$ . Example 3 will show how “difference equations” help to solve differential equations.

All these steps use the  $\lambda$ 's and the  $x$ 's. This section solves the constant coefficient problems that turn into linear algebra. It clarifies these simplest but most important differential equations—whose solution is completely based on growth factors  $e^{\lambda t}$ .

## Second Order Equations

**The most important equation in mechanics is  $my'' + by' + ky = 0$ .** The first term is the mass  $m$  times the acceleration  $a = y''$ . This term  $ma$  balances the force  $F$  (that is *Newton's Law*). The force includes the damping  $-by'$  and the elastic force  $-ky$ , proportional to distance moved. This is a second-order equation because it contains the second derivative  $y'' = d^2y/dt^2$ . It is still linear with constant coefficients  $m, b, k$ .

In a differential equations course, the method of solution is to substitute  $y = e^{\lambda t}$ . Each derivative of  $y$  brings down a factor  $\lambda$ . We want  $y = e^{\lambda t}$  to solve the equation:

$$m \frac{d^2y}{dt^2} + b \frac{dy}{dt} + ky = 0 \quad \text{becomes} \quad (m\lambda^2 + b\lambda + k)e^{\lambda t} = 0. \quad (8)$$

Everything depends on  $m\lambda^2 + b\lambda + k = 0$ . This equation for  $\lambda$  has two roots  $\lambda_1$  and  $\lambda_2$ . Then the equation for  $y$  has two pure solutions  $y_1 = e^{\lambda_1 t}$  and  $y_2 = e^{\lambda_2 t}$ . Their combinations  $c_1y_1 + c_2y_2$  give the complete solution unless  $\lambda_1 = \lambda_2$ .

In a linear algebra course we expect matrices and eigenvalues. Therefore we turn the scalar equation (with  $y''$ ) into a *vector equation for  $y$  and  $y'$* : first derivative only. Suppose the mass is  $m = 1$ . Two equations for  $\mathbf{u} = (y, y')$  give  $d\mathbf{u}/dt = A\mathbf{u}$ :

$$\begin{aligned} dy/dt &= y' \\ dy'/dt &= -ky - by' \end{aligned} \quad \text{converts to} \quad \frac{d}{dt} \begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -k & -b \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} = A\mathbf{u}. \quad (9)$$

The first equation  $dy/dt = y'$  is trivial (but true). The second is equation (8) connecting  $y''$  to  $y'$  and  $y$ . Together they connect  $\mathbf{u}'$  to  $\mathbf{u}$ . So we solve  $\mathbf{u}' = A\mathbf{u}$  by eigenvalues of  $A$ :

$$A - \lambda I = \begin{bmatrix} -\lambda & 1 \\ -k & -b - \lambda \end{bmatrix} \quad \text{has determinant} \quad \lambda^2 + b\lambda + k = 0.$$

**The equation for the  $\lambda$ 's is the same as (8)! It is still  $\lambda^2 + b\lambda + k = 0$ , since  $m = 1$ .** The roots  $\lambda_1$  and  $\lambda_2$  are now *eigenvalues of  $A$* . The eigenvectors and the solution are

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix} \quad \mathbf{u}(t) = c_1 e^{\lambda_1 t} \begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix} + c_2 e^{\lambda_2 t} \begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix}.$$

The first component of  $\mathbf{u}(t)$  has  $y = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$ —the same solution as before. It can't be anything else. In the second component of  $\mathbf{u}(t)$  you see the velocity  $dy/dt$ . The vector problem is completely consistent with the scalar problem. The 2 by 2 matrix  $A$  is called a *companion matrix*—a companion to the second order equation with  $y''$ .

**Example 3** *Motion around a circle with  $y'' + y = 0$  and  $y = \cos t$* 

This is our master equation with mass  $m = 1$  and stiffness  $k = 1$  and  $d = 0$ : no damping. Substitute  $y = e^{\lambda t}$  into  $y'' + y = 0$  to reach  $\lambda^2 + 1 = 0$ . The roots are  $\lambda = i$  and  $\lambda = -i$ . Then half of  $e^{it} + e^{-it}$  gives the solution  $y = \cos t$ .

As a first-order system, the initial values  $y(0) = 1, y'(0) = 0$  go into  $\mathbf{u}(0) = (1, 0)$ :

$$\text{Use } y'' = -y \quad \frac{d\mathbf{u}}{dt} = \frac{d}{dt} \begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} = A\mathbf{u}. \quad (10)$$

The eigenvalues of  $A$  are again the same  $\lambda = i$  and  $\lambda = -i$  (no surprise).  $A$  is anti-symmetric with eigenvectors  $\mathbf{x}_1 = (1, i)$  and  $\mathbf{x}_2 = (1, -i)$ . The combination that matches  $\mathbf{u}(0) = (1, 0)$  is  $\frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2)$ . Step 2 multiplies the  $x$ 's by  $e^{it}$  and  $e^{-it}$ . Step 3 combines the pure oscillations into  $\mathbf{u}(t)$  to find  $y = \cos t$  as expected:

$$\mathbf{u}(t) = \frac{1}{2}e^{it} \begin{bmatrix} 1 \\ i \end{bmatrix} + \frac{1}{2}e^{-it} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} \cos t \\ -\sin t \end{bmatrix}. \quad \text{This is } \begin{bmatrix} y(t) \\ y'(t) \end{bmatrix}.$$

All good. The vector  $\mathbf{u} = (\cos t, -\sin t)$  goes around a circle (Figure 6.3). The radius is 1 because  $\cos^2 t + \sin^2 t = 1$ .

**Difference Equations (optional)**

To display a circle on a screen, replace  $y'' = -y$  by a **difference equation**. Here are three choices using  $\mathbf{Y}(t + \Delta t) - 2\mathbf{Y}(t) + \mathbf{Y}(t - \Delta t)$ . Divide by  $(\Delta t)^2$  to approximate  $y''$ .

**F** Forward from  $n - 1$   
**C** Centered at time  $n$   
**B** Backward from  $n + 1$

$$\frac{Y_{n+1} - 2Y_n + Y_{n-1}}{(\Delta t)^2} = \begin{matrix} -Y_{n-1} & (11 \text{ F}) \\ -Y_n & (11 \text{ C}) \\ -Y_{n+1} & (11 \text{ B}) \end{matrix}$$

Figure 6.3 shows the exact  $y(t) = \cos t$  completing a circle at  $t = 2\pi$ . The three difference methods *don't* complete a perfect circle in 32 time steps of length  $\Delta t = 2\pi/32$ . Those pictures will be explained by eigenvalues:

**Forward  $|\lambda| > 1$  (spiral out)**   **Centered  $|\lambda| = 1$  (best)**   **Backward  $|\lambda| < 1$  (spiral in)**

The 2-step equations (11) reduce to 1-step systems  $\mathbf{U}_{n+1} = A\mathbf{U}_n$ . Instead of  $\mathbf{u} = (y, y')$  the discrete unknown is  $\mathbf{U}_n = (Y_n, Z_n)$ . We take  $n$  time steps  $\Delta t$  starting from  $\mathbf{U}_0$ :

$$\text{Forward (11F)} \quad \begin{matrix} Y_{n+1} = Y_n + \Delta t Z_n \\ Z_{n+1} = Z_n - \Delta t Y_n \end{matrix} \quad \text{becomes} \quad \mathbf{U}_{n+1} = \begin{bmatrix} 1 & \Delta t \\ -\Delta t & 1 \end{bmatrix} \begin{bmatrix} Y_n \\ Z_n \end{bmatrix} = A\mathbf{U}_n. \quad (12)$$

Those are like  $Y' = Z$  and  $Z' = -Y$ . They are first order equations involving times  $n$  and  $n + 1$ . Eliminating  $Z$  would bring back the “forward” second order equation (11 F).

My question is simple. *Do the points  $(Y_n, Z_n)$  stay on the circle  $Y^2 + Z^2 = 1$ ?* No, they are growing to infinity in Figure 6.3. **We are taking powers  $A^n$  and not  $e^{At}$ , so we test the magnitude  $|\lambda|$  and not the real parts of the eigenvalues.**

Eigenvalues of  $A$ 

$$\lambda = 1 \pm i\Delta t$$

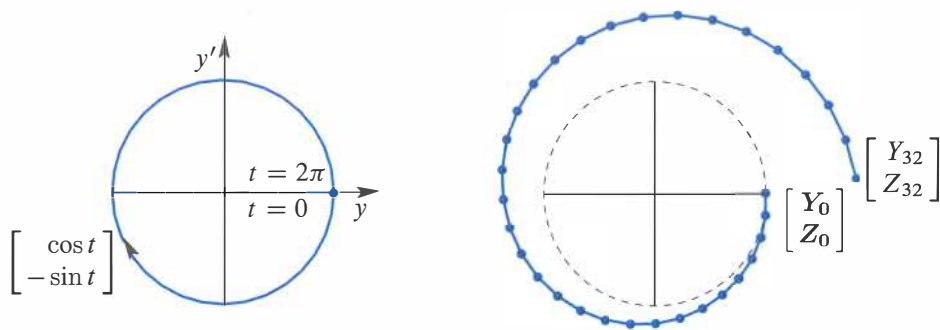
Then  $|\lambda| > 1$  and  $(Y_n, Z_n)$  spirals out

Figure 6.3: Exact  $u = (\cos t, -\sin t)$  on a circle. **Forward Euler spirals out** (32 steps).

The backward choice in (11 B) will do the opposite in Figure 6.4. Notice the new  $A$ :

$$\text{Backward} \quad \begin{aligned} Y_{n+1} &= Y_n + \Delta t Z_{n+1} \\ Z_{n+1} &= Z_n - \Delta t Y_{n+1} \end{aligned} \quad \text{is} \quad \begin{bmatrix} 1 & -\Delta t \\ \Delta t & 1 \end{bmatrix} \begin{bmatrix} Y_{n+1} \\ Z_{n+1} \end{bmatrix} = \begin{bmatrix} Y_n \\ Z_n \end{bmatrix} = U_n. \quad (13)$$

That matrix has eigenvalues  $1 \pm i\Delta t$ . But we *invert* it to reach  $U_{n+1}$  from  $U_n$ . Then  $|\lambda| < 1$  explains why *the solution spirals in* to  $(0,0)$  for backward differences.

On the right side of Figure 6.4 you see 32 steps with the *centered* choice. The solution stays close to the circle (Problem 28) if  $\Delta t < 2$ . This is the **leapfrog method**, constantly used. The second difference  $Y_{n+1} - 2Y_n + Y_{n-1}$  “leaps over” the center value  $Y_n$  in (11).

This is the way a chemist follows the motion of molecules (molecular dynamics leads to giant computations). Computational science is lively because one differential equation can be replaced by many difference equations—some unstable, some stable, some neutral. Problem 30 has a fourth (very good) method that stays right on the circle.

Real engineering and real physics deal with systems (not just a single mass at one point). The unknown  $y$  is a vector. The coefficient of  $y''$  is a *mass matrix*  $M$ , with  $n$  masses. The coefficient of  $y$  is a *stiffness matrix*  $K$ , not a number  $k$ . The coefficient of  $y'$  is a damping matrix which might be zero.

The vector equation  $My'' + Ky = f$  is a major part of computational mechanics. It is controlled by the eigenvalues of  $M^{-1}K$  in  $Kx = \lambda Mx$ .

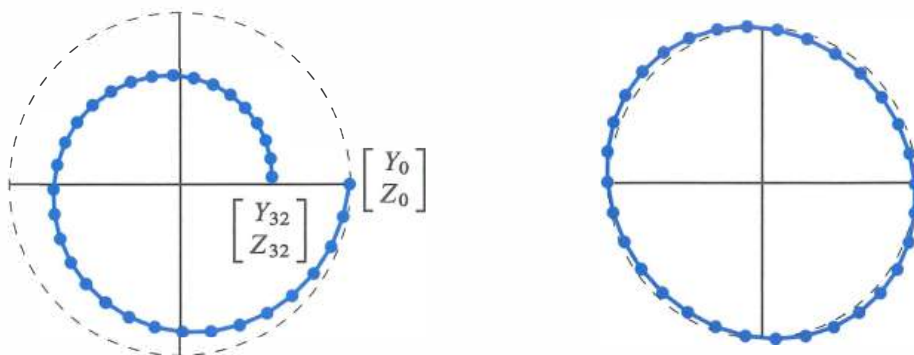


Figure 6.4: Backward differences spiral in. Leapfrog stays near the correct circle.

### Stability of 2 by 2 Matrices

For the solution of  $du/dt = Au$ , there is a fundamental question. *Does the solution approach  $u = 0$  as  $t \rightarrow \infty$ ?* Is the problem *stable*, by dissipating energy? A solution that includes  $e^t$  is unstable. Stability depends on the eigenvalues of  $A$ .

The complete solution  $u(t)$  is built from pure solutions  $e^{\lambda t}x$ . If the eigenvalue  $\lambda$  is real, we know exactly when  $e^{\lambda t}$  will approach zero: *The number  $\lambda$  must be negative.* If the eigenvalue is a complex number  $\lambda = r + is$ , *the real part  $r$  must be negative.* When  $e^{\lambda t}$  splits into  $e^{rt}e^{ist}$ , the factor  $e^{ist}$  has absolute value fixed at 1:

$$e^{ist} = \cos st + i \sin st \quad \text{has} \quad |e^{ist}|^2 = \cos^2 st + \sin^2 st = 1.$$

The real part of  $\lambda$  controls the growth ( $r > 0$ ) or the decay ( $r < 0$ ).

The question is: **Which matrices have negative eigenvalues?** More accurately, when are the **real parts of the  $\lambda$ 's all negative?** 2 by 2 matrices allow a clear answer.

**Stability**  $A$  is **stable** and  $u(t) \rightarrow 0$  when all eigenvalues  $\lambda$  have **negative real parts**.

The 2 by 2 matrix  $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$  must pass two tests:

$$\begin{aligned} \lambda_1 + \lambda_2 &< 0 \\ \lambda_1 \lambda_2 &> 0 \end{aligned}$$

The trace  $T = a + d$  must be negative.

The determinant  $D = ad - bc$  must be positive.

**Reason** If the  $\lambda$ 's are real and negative, their sum is negative. This is the trace  $T$ . Their product is positive. This is the determinant  $D$ . The argument also goes in the reverse direction. If  $D = \lambda_1 \lambda_2$  is positive, then  $\lambda_1$  and  $\lambda_2$  have the same sign. If  $T = \lambda_1 + \lambda_2$  is negative, that sign will be negative. We can test  $T$  and  $D$ .

If the  $\lambda$ 's are complex numbers, they must have the form  $r + is$  and  $r - is$ . Otherwise  $T$  and  $D$  will not be real. The determinant  $D$  is automatically positive, since  $(r + is)(r - is) = r^2 + s^2$ . The trace  $T$  is  $r + is + r - is = 2r$ . So a negative trace  $T$  means that the real part  $r$  is negative and the matrix is stable. Q.E.D.

Figure 6.5 shows the parabola  $T^2 = 4D$  separating real  $\lambda$ 's from complex  $\lambda$ 's. Solving  $\lambda^2 - T\lambda + D = 0$  involves the square root  $\sqrt{T^2 - 4D}$ . This is real below the parabola and imaginary above it. The stable region is the *upper left quarter* of the figure—where the trace  $T$  is negative and the determinant  $D$  is positive.

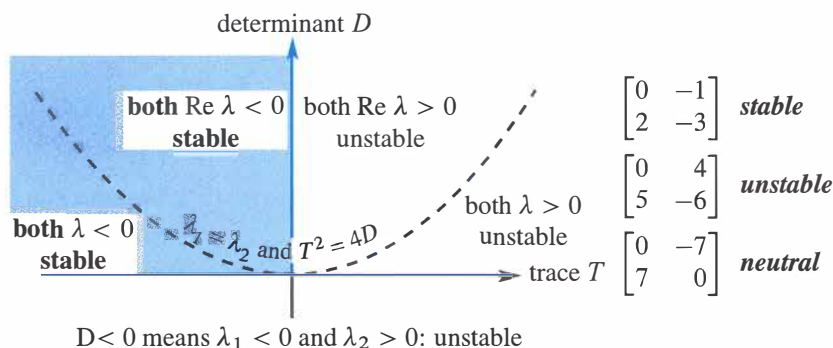


Figure 6.5: A 2 by 2 matrix is stable ( $\mathbf{u}(t) \rightarrow \mathbf{0}$ ) when **trace**  $< 0$  and **det**  $> 0$ .

## The Exponential of a Matrix

We want to write the solution  $\mathbf{u}(t)$  in a new form  $e^{At}\mathbf{u}(0)$ . First we have to say what  $e^{At}$  means, with a matrix in the exponent. To define  $e^{At}$  for matrices, we copy  $e^x$  for numbers.

The direct definition of  $e^x$  is by the infinite series  $1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots$ . When you change  $x$  to a square matrix  $At$ , this series defines the matrix exponential  $e^{At}$ :

**Matrix exponential**  $e^{At}$

$$e^{At} = I + At + \frac{1}{2}(At)^2 + \frac{1}{6}(At)^3 + \dots \quad (14)$$

**Its  $t$  derivative is**  $Ae^{At}$

$$A + A^2t + \frac{1}{2}A^3t^2 + \dots = Ae^{At}$$

**Its eigenvalues are**  $e^{\lambda t}$

$$(I + At + \frac{1}{2}(At)^2 + \dots)x = (1 + \lambda t + \frac{1}{2}(\lambda t)^2 + \dots)x$$

The number that divides  $(At)^n$  is “ $n$  factorial”. This is  $n! = (1)(2) \cdots (n-1)(n)$ . The factorials after 1, 2, 6 are  $4! = 24$  and  $5! = 120$ . They grow quickly. The series always converges and its derivative is always  $Ae^{At}$ . Therefore  $e^{At}\mathbf{u}(0)$  solves the differential equation with one quick formula—even if there is a shortage of eigenvectors.

I will use this series in Example 4, to see it work with a missing eigenvector. It will produce  $te^{\lambda t}$ . First let me reach  $Xe^{\lambda t}X^{-1}$  in the good (diagonalizable) case.

This chapter emphasizes how to find  $\mathbf{u}(t) = e^{At}\mathbf{u}(0)$  by diagonalization. Assume  $A$  does have  $n$  independent eigenvectors, so it is diagonalizable. Substitute  $A = X\Lambda X^{-1}$  into the series for  $e^{At}$ . Whenever  $X\Lambda X^{-1}X\Lambda X^{-1}$  appears, cancel  $X^{-1}X$  in the middle:

**Use the series** 
$$e^{At} = I + X\Lambda X^{-1}t + \frac{1}{2}(X\Lambda X^{-1}t)(X\Lambda X^{-1}t) + \dots$$

**Factor out  $X$  and  $X^{-1}$**  
$$= X \left[ I + \Lambda t + \frac{1}{2}(\Lambda t)^2 + \dots \right] X^{-1} \quad (15)$$

**$e^{At}$  is diagonalized!** 
$$e^{At} = X e^{\Lambda t} X^{-1}.$$

$e^{At}$  has the same eigenvector matrix  $X$  as  $A$ . Then  $\Lambda$  is a diagonal matrix and so is  $e^{\Lambda t}$ . The numbers  $e^{\lambda_i t}$  are on the diagonal. Multiply  $X e^{\Lambda t} X^{-1} \mathbf{u}(0)$  to recognize  $\mathbf{u}(t)$ :

$$e^{At}\mathbf{u}(0) = X e^{\Lambda t} X^{-1} \mathbf{u}(0) = \begin{bmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}. \quad (16)$$

This solution  $e^{At}\mathbf{u}(0)$  is the same answer that came in equation (6) from three steps:

1.  $\mathbf{u}(0) = c_1 \mathbf{x}_1 + \cdots + c_n \mathbf{x}_n = X\mathbf{c}$ . Here we need  $n$  independent eigenvectors.
2. Multiply each  $\mathbf{x}_i$  by its growth factor  $e^{\lambda_i t}$  to follow it forward in time.
3. The best form of  $e^{At}\mathbf{u}(0)$  is  $\mathbf{u}(t) = c_1 e^{\lambda_1 t} \mathbf{x}_1 + \cdots + c_n e^{\lambda_n t} \mathbf{x}_n$ . (17)

**Example 4** When you substitute  $y = e^{\lambda t}$  into  $y'' - 2y' + y = 0$ , you get an equation with **repeated roots**:  $\lambda^2 - 2\lambda + 1 = 0$  is  $(\lambda - 1)^2 = 0$  with  $\lambda = 1, 1$ . A differential equations course would propose  $e^t$  and  $te^t$  as two independent solutions. Here we discover why.

Linear algebra reduces  $y'' - 2y' + y = 0$  to a vector equation for  $\mathbf{u} = (y, y')$ :

$$\frac{d}{dt} \begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} y' \\ 2y' - y \end{bmatrix} \quad \text{is} \quad \frac{d\mathbf{u}}{dt} = A\mathbf{u} = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \mathbf{u}. \quad (18)$$

$A$  has a **repeated eigenvalue**  $\lambda = 1, 1$  (with trace = 2 and  $\det A = 1$ ). The only eigenvectors are multiples of  $\mathbf{x} = (1, 1)$ . *Diagonalization is not possible*,  $A$  has only one line of eigenvectors. So we compute  $e^{At}$  from its definition as a series:

**Short series** 
$$e^{At} = e^{It} e^{(A-I)t} = e^t [I + (A - I)t]. \quad (19)$$

That “infinite” series for  $e^{(A-I)t}$  ended quickly because  $(A - I)^2$  is the zero matrix! You can see  $te^t$  in equation (19). The first component of  $e^{At}\mathbf{u}(0)$  is our answer  $y(t)$ :

$$\begin{bmatrix} y \\ y' \end{bmatrix} = e^t \left[ I + \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} t \right] \begin{bmatrix} y(0) \\ y'(0) \end{bmatrix} \quad y(t) = e^t y(0) - te^t y(0) + te^t y'(0).$$

**Example 5** Use the infinite series to find  $e^{At}$  for  $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ . Notice that  $A^4 = I$ :

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad A^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \quad A^3 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad A^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$A^5, A^6, A^7, A^8$  will be a repeat of  $A, A^2, A^3, A^4$ . The top right corner has 1, 0, -1, 0 repeating over and over in powers of  $A$ . Then  $t - \frac{1}{6}t^3$  starts the infinite series for  $e^{At}$  in that top right corner, and  $1 - \frac{1}{2}t^2$  starts the top left corner:

$$e^{At} = I + At + \frac{1}{2}(At)^2 + \frac{1}{6}(At)^3 + \cdots = \begin{bmatrix} 1 - \frac{1}{2}t^2 + \cdots & t - \frac{1}{6}t^3 + \cdots \\ -t + \frac{1}{6}t^3 - \cdots & 1 - \frac{1}{2}t^2 + \cdots \end{bmatrix}.$$

The top row of that matrix  $e^{At}$  shows the infinite series for the cosine and sine!

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad e^{At} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}. \quad (20)$$

$A$  is an antisymmetric matrix ( $A^T = -A$ ). Its exponential  $e^{At}$  is an orthogonal matrix. The eigenvalues of  $A$  are  $i$  and  $-i$ . The eigenvalues of  $e^{At}$  are  $e^{it}$  and  $e^{-it}$ . Three rules:

- 1  $e^{At}$  always has the inverse  $e^{-At}$ .
- 2 The eigenvalues of  $e^{At}$  are always  $e^{\lambda t}$ .
- 3 When  $A$  is antisymmetric,  $e^{At}$  is orthogonal. Inverse = transpose =  $e^{-At}$ .

Antisymmetric is the same as “skew-symmetric”. Those matrices have pure imaginary eigenvalues like  $i$  and  $-i$ . Then  $e^{At}$  has eigenvalues like  $e^{it}$  and  $e^{-it}$ . Their absolute value is 1: neutral stability, pure oscillation, energy is conserved. So  $\|u(t)\| = \|u(0)\|$ .

Our final example has a triangular matrix  $A$ . Then the eigenvector matrix  $X$  is triangular. So are  $X^{-1}$  and  $e^{At}$ . You will see the two forms of the solution: a combination of eigenvectors and the short form  $e^{At}u(0)$ .

**Example 6** Solve  $\frac{du}{dt} = Au = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} u$  starting from  $u(0) = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$  at  $t = 0$ .

**Solution** The eigenvalues 1 and 2 are on the diagonal of  $A$  (since  $A$  is triangular). The eigenvectors are  $(1, 0)$  and  $(1, 1)$ . The starting  $u(0)$  is  $x_1 + x_2$  so  $c_1 = c_2 = 1$ . Then  $u(t)$  is the same combination of pure exponentials (no  $te^{\lambda t}$  when  $\lambda = 1$  and 2):

$$\text{Solution to } u' = Au \quad u(t) = e^t \begin{bmatrix} 1 \\ 0 \end{bmatrix} + e^{2t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

That is the clearest form. But the matrix form with  $e^{At}$  produces  $u(t)$  for every  $u(0)$ :

$$u(t) = X e^{At} X^{-1} u(0) \text{ is } \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e^t & \\ & e^{2t} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} u(0) = \begin{bmatrix} e^t & e^{2t} - e^t \\ 0 & e^{2t} \end{bmatrix} u(0).$$

**That last matrix is  $e^{At}$ .** It is nice because  $A$  is triangular. The situation is the same as for  $Ax = b$  and inverses. We don't need  $A^{-1}$  to find  $x$ , and we don't need  $e^{At}$  to solve  $du/dt = Au$ . But as quick formulas for the answers,  $A^{-1}b$  and  $e^{At}u(0)$  are unbeatable.

### ■ REVIEW OF THE KEY IDEAS ■

1. The equation  $\mathbf{u}' = A\mathbf{u}$  is linear with constant coefficients in  $A$ . Start from  $\mathbf{u}(0)$ .
2. Its solution is usually a combination of exponentials, involving every  $\lambda$  and  $\mathbf{x}$ :

**Independent eigenvectors**       $\mathbf{u}(t) = c_1 e^{\lambda_1 t} \mathbf{x}_1 + \cdots + c_n e^{\lambda_n t} \mathbf{x}_n.$

3. The constants  $c_1, \dots, c_n$  are determined by  $\mathbf{u}(0) = c_1 \mathbf{x}_1 + \cdots + c_n \mathbf{x}_n = X\mathbf{c}$ .
4.  $\mathbf{u}(t)$  approaches zero (**stability**) if every  $\lambda$  has negative real part: All  $e^{\lambda t} \rightarrow 0$ .
5. Solutions have the short form  $\mathbf{u}(t) = e^{At} \mathbf{u}(0)$ , with the matrix exponential  $e^{At}$ .
6. Equations with  $y''$  reduce to  $\mathbf{u}' = A\mathbf{u}$  by combining  $y$  and  $y'$  into the vector  $\mathbf{u}$ .

### ■ WORKED EXAMPLES ■

**6.3 A** Solve  $y'' + 4y' + 3y = 0$  by substituting  $e^{\lambda t}$  and also by linear algebra.

**Solution** Substituting  $y = e^{\lambda t}$  yields  $(\lambda^2 + 4\lambda + 3)e^{\lambda t} = 0$ . That quadratic factors into  $\lambda^2 + 4\lambda + 3 = (\lambda + 1)(\lambda + 3) = 0$ . Therefore  $\lambda_1 = -1$  and  $\lambda_2 = -3$ . The pure solutions are  $y_1 = e^{-t}$  and  $y_2 = e^{-3t}$ . The complete solution  $y = c_1 y_1 + c_2 y_2$  approaches zero.

To use linear algebra we set  $\mathbf{u} = (y, y')$ . Then the vector equation is  $\mathbf{u}' = A\mathbf{u}$ :

$$\begin{aligned} dy/dt &= y' \\ dy'/dt &= -3y - 4y' \end{aligned} \quad \text{converts to} \quad \frac{d\mathbf{u}}{dt} = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix} \mathbf{u}.$$

This  $A$  is a “companion matrix” and its eigenvalues are again  $-1$  and  $-3$ :

**Same quadratic**       $\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -3 & -4 - \lambda \end{vmatrix} = \lambda^2 + 4\lambda + 3 = 0.$

The eigenvectors of  $A$  are  $(1, \lambda_1)$  and  $(1, \lambda_2)$ . Either way, the decay in  $y(t)$  comes from  $e^{-t}$  and  $e^{-3t}$ . With constant coefficients, calculus leads to linear algebra  $A\mathbf{x} = \lambda\mathbf{x}$ .

**Note** In linear algebra the serious danger is a shortage of eigenvectors. Our eigenvectors  $(1, \lambda_1)$  and  $(1, \lambda_2)$  are the same if  $\lambda_1 = \lambda_2$ . Then we can't diagonalize  $A$ . In this case we don't yet have two independent solutions to  $d\mathbf{u}/dt = A\mathbf{u}$ .

In differential equations the danger is also a repeated  $\lambda$ . After  $y = e^{\lambda t}$ , a second solution has to be found. It turns out to be  $y = te^{\lambda t}$ . This “impure” solution (with an extra  $t$ ) appears in the matrix exponential  $e^{At}$ . Example 4 showed how.

**6.3 B** Find the eigenvalues and eigenvectors of  $A$ . Then write  $\mathbf{u}(0) = (0, 2\sqrt{2}, 0)$  as a combination of the eigenvectors. Solve both equations  $\mathbf{u}' = A\mathbf{u}$  and  $\mathbf{u}'' = A\mathbf{u}$ :

$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \mathbf{u} \quad \text{and} \quad \frac{d^2\mathbf{u}}{dt^2} = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \mathbf{u} \quad \text{with} \quad \frac{d\mathbf{u}}{dt}(0) = \mathbf{0}.$$

$\mathbf{u}' = A\mathbf{u}$  is like the heat equation  $\partial u / \partial t = \partial^2 u / \partial x^2$ .

Its solution  $u(t)$  will decay ( $A$  has negative eigenvalues).

$\mathbf{u}'' = A\mathbf{u}$  is like the wave equation  $\partial^2 u / \partial t^2 = \partial^2 u / \partial x^2$ .

Its solution will oscillate (the square roots of  $\lambda$  are imaginary).

**Solution** The eigenvalues and eigenvectors come from  $\det(A - \lambda I) = 0$ :

$$\det(A - \lambda I) = \begin{vmatrix} -2 - \lambda & 1 & 0 \\ 1 & -2 - \lambda & 1 \\ 0 & 1 & -2 - \lambda \end{vmatrix} = (-2 - \lambda)[(-2 - \lambda)^2 - 2] = 0.$$

One eigenvalue is  $\lambda = -2$ , when  $-2 - \lambda$  is zero. The other factor is  $\lambda^2 + 4\lambda + 2$ , so the other eigenvalues (also real and negative) are  $\lambda = -2 \pm \sqrt{2}$ . Find the eigenvectors:

$$\lambda = -2 \quad (A + 2I)\mathbf{x} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{for } \mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$\lambda = -2 - \sqrt{2} \quad (A - \lambda I)\mathbf{x} = \begin{bmatrix} \sqrt{2} & 1 & 0 \\ 1 & \sqrt{2} & 1 \\ 0 & 1 & \sqrt{2} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{for } \mathbf{x}_2 = \begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix}$$

$$\lambda = -2 + \sqrt{2} \quad (A - \lambda I)\mathbf{x} = \begin{bmatrix} -\sqrt{2} & 1 & 0 \\ 1 & -\sqrt{2} & 1 \\ 0 & 1 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{for } \mathbf{x}_3 = \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}$$

The eigenvectors are *orthogonal* (proved in Section 6.4 for all symmetric matrices). All three  $\lambda_i$  are negative. This  $A$  is *negative definite* and  $e^{At}$  decays to zero (stability).

The starting  $\mathbf{u}(0) = (0, 2\sqrt{2}, 0)$  is  $\mathbf{x}_3 - \mathbf{x}_2$ . The solution is  $\mathbf{u}(t) = e^{\lambda_3 t} \mathbf{x}_3 - e^{\lambda_2 t} \mathbf{x}_2$ .

**Heat equation** In Figure 6.6a, the temperature at the center starts at  $2\sqrt{2}$ . Heat diffuses into the neighboring boxes and then to the outside boxes (frozen at  $0^\circ$ ). The rate of heat flow between boxes is the temperature difference. From box 2, heat flows left and right at the rate  $u_1 - u_2$  and  $u_3 - u_2$ . So the flow out is  $u_1 - 2u_2 + u_3$  in the second row of  $A\mathbf{u}$ .

**Wave equation**  $d^2\mathbf{u}/dt^2 = A\mathbf{u}$  has the same eigenvectors  $\mathbf{x}$ . But now the eigenvalues  $\lambda$  lead to **oscillations**  $e^{i\omega t}\mathbf{x}$  and  $e^{-i\omega t}\mathbf{x}$ . The frequencies come from  $\omega^2 = -\lambda$ :

$$\frac{d^2}{dt^2}(e^{i\omega t}\mathbf{x}) = A(e^{i\omega t}\mathbf{x}) \quad \text{becomes} \quad (i\omega)^2 e^{i\omega t}\mathbf{x} = \lambda e^{i\omega t}\mathbf{x} \quad \text{and} \quad \omega^2 = -\lambda.$$

There are two square roots of  $-\lambda$ , so we have  $e^{i\omega t}\mathbf{x}$  and  $e^{-i\omega t}\mathbf{x}$ . With three eigenvectors this makes *six* solutions to  $\mathbf{u}'' = A\mathbf{u}$ . A combination will match the six components of  $\mathbf{u}(0)$  and  $\mathbf{u}'(0)$ . Since  $\mathbf{u}' = \mathbf{0}$  in this problem,  $e^{i\omega t}\mathbf{x}$  and  $e^{-i\omega t}\mathbf{x}$  produce  $2\cos\omega t\mathbf{x}$ .

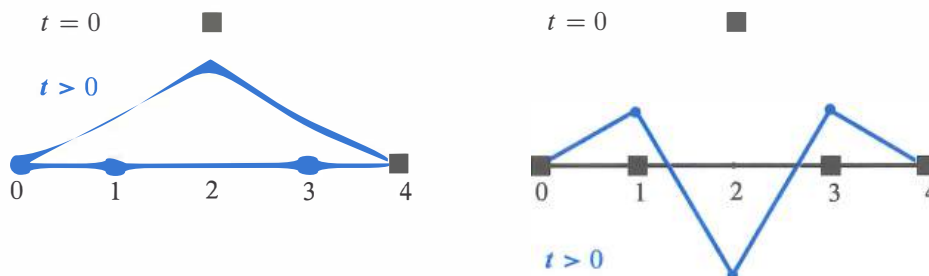


Figure 6.6: Heat diffuses away from box 2 (left). Wave travels from box 2 (right).

**6.3 C** Solve the four equations  $da/dt = 0, db/dt = a, dc/dt = 2b, dz/dt = 3c$  in that order starting from  $\mathbf{u}(0) = (a(0), b(0), c(0), z(0))$ . Solve the same equations by the matrix exponential in  $\mathbf{u}(t) = e^{At}\mathbf{u}(0)$ .

Four equations  
 $\lambda = 0, 0, 0, 0$   
 Eigenvalues on the diagonal

$$\frac{d}{dt} \begin{bmatrix} a \\ b \\ c \\ z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ z \end{bmatrix} \quad \text{is} \quad \frac{d\mathbf{u}}{dt} = A\mathbf{u}.$$

First find  $A^2, A^3, A^4$  and  $e^{At} = I + At + \frac{1}{2}(At)^2 + \frac{1}{6}(At)^3$ . Why does the series stop? Why is it true that  $(e^A)(e^A) = (e^{2A})$ ? **Always**  $e^{As}$  times  $e^{At}$  is  $e^{A(s+t)}$ .

**Solution 1** Integrate  $da/dt = 0$ , then  $db/dt = a$ , then  $dc/dt = 2b$  and  $dz/dt = 3c$ :

$$\begin{aligned} a(t) &= a(0) \\ b(t) &= ta(0) + b(0) \\ c(t) &= t^2a(0) + 2tb(0) + c(0) \\ z(t) &= t^3a(0) + 3t^2b(0) + 3tc(0) + z(0) \end{aligned}$$

The 4 by 4 matrix which is multiplying  $a(0), b(0), c(0), d(0)$  to produce  $a(t), b(t), c(t), d(t)$  must be the same  $e^{At}$  as below

**Solution 2** The powers of  $A$  (strictly triangular) are all zero after  $A^3$ .

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix} \quad A^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \end{bmatrix} \quad A^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \end{bmatrix} \quad A^4 = 0$$

The diagonals move down at each step. So the series for  $e^{At}$  stops after four terms:

Same  $e^{At}$  as in Solution 1

$$e^{At} = I + At + \frac{(At)^2}{2} + \frac{(At)^3}{6} = \begin{bmatrix} 1 & & & \\ t & 1 & & \\ t^2 & 2t & 1 & \\ t^3 & 3t^2 & 3t & 1 \end{bmatrix}$$

The square of  $e^A$  is  $e^{2A}$ . But  $e^A e^B$  and  $e^B e^A$  and  $e^{A+B}$  can be all different.

## Problem Set 6.3

- 1 Find two  $\lambda$ 's and  $\mathbf{x}$ 's so that  $\mathbf{u} = e^{\lambda t} \mathbf{x}$  solves

$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} 4 & 3 \\ 0 & 1 \end{bmatrix} \mathbf{u}.$$

What combination  $\mathbf{u} = c_1 e^{\lambda_1 t} \mathbf{x}_1 + c_2 e^{\lambda_2 t} \mathbf{x}_2$  starts from  $\mathbf{u}(0) = (5, -2)$ ?

- 2 Solve Problem 1 for  $\mathbf{u} = (y, z)$  by back substitution,  $z$  before  $y$ :

$$\text{Solve } \frac{dz}{dt} = z \text{ from } z(0) = -2. \quad \text{Then solve } \frac{dy}{dt} = 4y + 3z \text{ from } y(0) = 5.$$

The solution for  $y$  will be a combination of  $e^{4t}$  and  $e^t$ . The  $\lambda$ 's are 4 and 1.

- 3 (a) If every column of  $A$  adds to zero, why is  $\lambda = 0$  an eigenvalue?  
 (b) With negative diagonal and positive off-diagonal adding to zero,  $\mathbf{u}' = A\mathbf{u}$  will be a "continuous" Markov equation. Find the eigenvalues and eigenvectors, and the *steady state* as  $t \rightarrow \infty$

$$\text{Solve } \frac{d\mathbf{u}}{dt} = \begin{bmatrix} -2 & 3 \\ 2 & -3 \end{bmatrix} \mathbf{u} \text{ with } \mathbf{u}(0) = \begin{bmatrix} 4 \\ 1 \end{bmatrix}. \text{ What is } \mathbf{u}(\infty)?$$

- 4 A door is opened between rooms that hold  $v(0) = 30$  people and  $w(0) = 10$  people. The movement between rooms is proportional to the difference  $v - w$ :

$$\frac{dv}{dt} = w - v \quad \text{and} \quad \frac{dw}{dt} = v - w.$$

Show that the total  $v + w$  is constant (40 people). Find the matrix in  $d\mathbf{u}/dt = A\mathbf{u}$  and its eigenvalues and eigenvectors. What are  $v$  and  $w$  at  $t = 1$  and  $t = \infty$ ?

- 5 Reverse the diffusion of people in Problem 4 to  $d\mathbf{u}/dt = -A\mathbf{u}$ :

$$\frac{dv}{dt} = v - w \quad \text{and} \quad \frac{dw}{dt} = w - v.$$

The total  $v + w$  still remains constant. How are the  $\lambda$ 's changed now that  $A$  is changed to  $-A$ ? But show that  $v(t)$  grows to infinity from  $v(0) = 30$ .

- 6  $A$  has real eigenvalues but  $B$  has complex eigenvalues:

$$A = \begin{bmatrix} a & 1 \\ 1 & a \end{bmatrix} \quad B = \begin{bmatrix} b & -1 \\ 1 & b \end{bmatrix} \quad (a \text{ and } b \text{ are real})$$

Find the conditions on  $a$  and  $b$  so that all solutions of  $d\mathbf{u}/dt = A\mathbf{u}$  and  $d\mathbf{v}/dt = B\mathbf{v}$  approach zero as  $t \rightarrow \infty$ :  $\operatorname{Re} \lambda < 0$  for all eigenvalues.

- 7 Suppose  $P$  is the projection matrix onto the  $45^\circ$  line  $y = x$  in  $\mathbf{R}^2$ . What are its eigenvalues? If  $du/dt = -Pu$  (notice minus sign) can you find the limit of  $u(t)$  at  $t = \infty$  starting from  $u(0) = (3, 1)$ ?
- 8 The rabbit population shows fast growth (from  $6r$ ) but loss to wolves (from  $-2w$ ). The wolf population always grows in this model ( $-w^2$  would control wolves):

$$\frac{dr}{dt} = 6r - 2w \quad \text{and} \quad \frac{dw}{dt} = 2r + w.$$

Find the eigenvalues and eigenvectors. If  $r(0) = w(0) = 30$  what are the populations at time  $t$ ? After a long time, what is the ratio of rabbits to wolves?

- 9 (a) Write  $(4, 0)$  as a combination  $c_1x_1 + c_2x_2$  of these two eigenvectors of  $A$ :

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = i \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = -i \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

- (b) The solution to  $du/dt = Au$  starting from  $(4, 0)$  is  $c_1e^{it}x_1 + c_2e^{-it}x_2$ . Substitute  $e^{it} = \cos t + i \sin t$  and  $e^{-it} = \cos t - i \sin t$  to find  $u(t)$ .

**Questions 10–13 reduce second-order equations to first-order systems for  $(y, y')$ .**

- 10 Find  $A$  to change the scalar equation  $y'' = 5y' + 4y$  into a vector equation for  $u = (y, y')$ :

$$\frac{du}{dt} = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} & \\ & \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} = Au.$$

What are the eigenvalues of  $A$ ? Find them also by substituting  $y = e^{\lambda t}$  into  $y'' = 5y' + 4y$ .

- 11 The solution to  $y'' = 0$  is a straight line  $y = C + Dt$ . Convert to a matrix equation:

$$\frac{d}{dt} \begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} \text{ has the solution } \begin{bmatrix} y \\ y' \end{bmatrix} = e^{At} \begin{bmatrix} y(0) \\ y'(0) \end{bmatrix}.$$

This matrix  $A$  has  $\lambda = 0, 0$  and it cannot be diagonalized. Find  $A^2$  and compute  $e^{At} = I + At + \frac{1}{2}A^2t^2 + \dots$ . Multiply your  $e^{At}$  times  $(y(0), y'(0))$  to check the straight line  $y(t) = y(0) + y'(0)t$ .

- 12 Substitute  $y = e^{\lambda t}$  into  $y'' = 6y' - 9y$  to show that  $\lambda = 3$  is a repeated root. This is trouble; we need a second solution after  $e^{3t}$ . The matrix equation is

$$\frac{d}{dt} \begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -9 & 6 \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix}.$$

Show that this matrix has  $\lambda = 3, 3$  and only one line of eigenvectors. *Trouble here too.* Show that the second solution to  $y'' = 6y' - 9y$  is  $y = te^{3t}$ .

- 13 (a) Write down two familiar functions that solve the equation  $d^2y/dt^2 = -9y$ . Which one starts with  $y(0) = 3$  and  $y'(0) = 0$ ?
- (b) This second-order equation  $y'' = -9y$  produces a vector equation  $\mathbf{u}' = A\mathbf{u}$ :

$$\mathbf{u} = \begin{bmatrix} y \\ y' \end{bmatrix} \quad \frac{d\mathbf{u}}{dt} = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -9 & 0 \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} = A\mathbf{u}.$$

Find  $\mathbf{u}(t)$  by using the eigenvalues and eigenvectors of  $A$ :  $\mathbf{u}(0) = (3, 0)$ .

- 14 The matrix in this question is skew-symmetric ( $A^T = -A$ ):

$$\frac{d\mathbf{u}}{dt} = \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix} \mathbf{u} \quad \text{or} \quad \begin{aligned} u'_1 &= cu_2 - bu_3 \\ u'_2 &= au_3 - cu_1 \\ u'_3 &= bu_1 - au_2. \end{aligned}$$

- (a) The derivative of  $\|\mathbf{u}(t)\|^2 = u_1^2 + u_2^2 + u_3^2$  is  $2u_1u'_1 + 2u_2u'_2 + 2u_3u'_3$ . Substitute  $u'_1, u'_2, u'_3$  to get zero. Then  $\|\mathbf{u}(t)\|^2$  stays equal to  $\|\mathbf{u}(0)\|^2$ .
- (b) When  $A$  is skew-symmetric,  $Q = e^{At}$  is orthogonal. Prove  $Q^T = e^{-At}$  from the series for  $Q = e^{At}$ . Then  $Q^T Q = I$ .
- 15 A particular solution to  $d\mathbf{u}/dt = A\mathbf{u} - \mathbf{b}$  is  $\mathbf{u}_p = A^{-1}\mathbf{b}$ , if  $A$  is invertible. The usual solutions to  $d\mathbf{u}/dt = A\mathbf{u}$  give  $\mathbf{u}_n$ . Find the complete solution  $\mathbf{u} = \mathbf{u}_p + \mathbf{u}_n$ :
- (a)  $\frac{du}{dt} = u - 4$       (b)  $\frac{d\mathbf{u}}{dt} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \mathbf{u} - \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ .
- 16 If  $c$  is not an eigenvalue of  $A$ , substitute  $\mathbf{u} = e^{ct}\mathbf{v}$  and find a particular solution to  $d\mathbf{u}/dt = A\mathbf{u} - e^{ct}\mathbf{b}$ . How does it break down when  $c$  is an eigenvalue of  $A$ ? The “nullspace” of  $d\mathbf{u}/dt = A\mathbf{u}$  contains the usual solutions  $e^{\lambda_i t}\mathbf{x}_i$ .
- 17 Find a matrix  $A$  to illustrate each of the unstable regions in Figure 6.5:
- (a)  $\lambda_1 < 0$  and  $\lambda_2 > 0$     (b)  $\lambda_1 > 0$  and  $\lambda_2 > 0$     (c)  $\lambda = a \pm ib$  with  $a > 0$ .

**Questions 18–27 are about the matrix exponential  $e^{At}$ .**

- 18 Write five terms of the infinite series for  $e^{At}$ . Take the  $t$  derivative of each term. Show that you have four terms of  $Ae^{At}$ . Conclusion:  $e^{At}\mathbf{u}_0$  solves  $\mathbf{u}' = A\mathbf{u}$ .
- 19 The matrix  $B = \begin{bmatrix} 0 & -4 \\ 0 & 0 \end{bmatrix}$  has  $B^2 = 0$ . Find  $e^{Bt}$  from a (short) infinite series. Check that the derivative of  $e^{Bt}$  is  $Be^{Bt}$ .
- 20 Starting from  $\mathbf{u}(0)$  the solution at time  $T$  is  $e^{AT}\mathbf{u}(0)$ . Go an additional time  $t$  to reach  $e^{At}e^{AT}\mathbf{u}(0)$ . This solution at time  $t + T$  can also be written as \_\_\_\_\_. Conclusion:  $e^{At}$  times  $e^{AT}$  equals \_\_\_\_\_.
- 21 Write  $A = \begin{bmatrix} 1 & 4 \\ 0 & 0 \end{bmatrix}$  in the form  $X\Lambda X^{-1}$ . Find  $e^{At}$  from  $Xe^{\Lambda t}X^{-1}$ .

- 22** If  $A^2 = A$  show that the infinite series produces  $e^{At} = I + (e^t - 1)A$ . For  $A = \begin{bmatrix} 1 & 4 \\ 0 & 0 \end{bmatrix}$  in Problem 21 this gives  $e^{At} = \underline{\hspace{2cm}}$ .
- 23** Generally  $e^A e^B$  is different from  $e^B e^A$ . They are both different from  $e^{A+B}$ . Check this using Problems 21–22 and 19. (If  $AB = BA$ , all three are the same.)

$$A = \begin{bmatrix} 1 & 4 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & -4 \\ 0 & 0 \end{bmatrix} \quad A + B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

- 24** Write  $A = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix}$  as  $X\Lambda X^{-1}$ . Multiply  $Xe^{\Lambda t}X^{-1}$  to find the matrix exponential  $e^{At}$ . Check  $e^{At}$  and the derivative of  $e^{At}$  when  $t = 0$ .
- 25** Put  $A = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$  into the infinite series to find  $e^{At}$ . First compute  $A^2$  and  $A^n$ :

$$e^{At} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} t & 3t \\ 0 & 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix} + \cdots = \begin{bmatrix} e^t & \quad \\ 0 & \quad \end{bmatrix}.$$

- 26** (Recommended) Give two reasons why the matrix exponential  $e^{At}$  is never singular:
- (a) Write down its inverse.
- (b) Why are these eigenvalues nonzero? If  $Ax = \lambda x$  then  $e^{At}x = \underline{\hspace{2cm}}x$ .
- 27** Find a solution  $x(t), y(t)$  that gets large as  $t \rightarrow \infty$ . To avoid this instability a scientist exchanged the two equations:

$$\begin{array}{ll} dx/dt = 0x - 4y & \\ dy/dt = -2x + 2y & \end{array} \quad \text{becomes} \quad \begin{array}{ll} dy/dt = -2x + 2y & \\ dx/dt = 0x - 4y. & \end{array}$$

Now the matrix  $\begin{bmatrix} -2 & 2 \\ 0 & -4 \end{bmatrix}$  is stable. It has negative eigenvalues. How can this be?

### Challenge Problems

- 28** Centering  $y'' = -y$  in Example 3 will produce  $Y_{n+1} - 2Y_n + Y_{n-1} = -(\Delta t)^2 Y_n$ . This can be written as a one-step difference equation for  $U = (Y, Z)$ :

$$\begin{array}{l} Y_{n+1} = Y_n + \Delta t Z_n \\ Z_{n+1} = Z_n - \Delta t Y_{n+1} \end{array} \quad \begin{bmatrix} 1 & 0 \\ \Delta t & 1 \end{bmatrix} \begin{bmatrix} Y_{n+1} \\ Z_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} Y_n \\ Z_n \end{bmatrix}$$

Invert the matrix on the left side to write this as  $U_{n+1} = AU_n$ . Show that  $\det A = 1$ . Choose the large time step  $\Delta t = 1$  and find the eigenvalues  $\lambda_1$  and  $\lambda_2 = \bar{\lambda}_1$  of  $A$ :

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} \text{ has } |\lambda_1| = |\lambda_2| = 1. \text{ Show that } A^6 \text{ is exactly } I.$$

After 6 steps to  $t = 6$ ,  $U_6$  equals  $U_0$ . The exact  $y = \cos t$  returns to 1 at  $t = 2\pi$ .

- 29 That centered choice (*leapfrog method*) in Problem 28 is very successful for small time steps  $\Delta t$ . But find the eigenvalues of  $A$  for  $\Delta t = \sqrt{2}$  and 2:

$$A = \begin{bmatrix} 1 & \sqrt{2} \\ -\sqrt{2} & -1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 2 \\ -2 & -3 \end{bmatrix}.$$

Both matrices have  $|\lambda| = 1$ . Compute  $A^4$  in both cases and find the eigenvectors of  $A$ . That second value  $\Delta t = 2$  is at the border of instability. Any time step  $\Delta t > 2$  will lead to  $|\lambda| > 1$ , and the powers in  $U_n = A^n U_0$  will explode.

*Note* You might say that nobody would compute with  $\Delta t > 2$ . But if an atom vibrates with  $y'' = -1000000y$ , then  $\Delta t > .0002$  will give instability. Leapfrog has a very strict stability limit.  $Y_{n+1} = Y_n + 3Z_n$  and  $Z_{n+1} = Z_n - 3Y_{n+1}$  will explode because  $\Delta t = 3$  is too large. The matrix has  $|\lambda| > 1$ .

- 30 Another good idea for  $y'' = -y$  is the trapezoidal method (half forward/half back). *This may be the best way to keep  $(Y_n, Z_n)$  exactly on a circle.*

$$\text{Trapezoidal} \quad \begin{bmatrix} 1 & -\Delta t/2 \\ \Delta t/2 & 1 \end{bmatrix} \begin{bmatrix} Y_{n+1} \\ Z_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t/2 \\ -\Delta t/2 & 1 \end{bmatrix} \begin{bmatrix} Y_n \\ Z_n \end{bmatrix}.$$

- (a) Invert the left matrix to write this equation as  $U_{n+1} = AU_n$ . Show that  $A$  is an orthogonal matrix:  $A^T A = I$ . **These points  $U_n$  never leave the circle.**  $A = (I - B)^{-1}(I + B)$  is always an orthogonal matrix if  $B^T = -B$ .
- (b) (Optional MATLAB) Take 32 steps from  $U_0 = (1, 0)$  to  $U_{32}$  with  $\Delta t = 2\pi/32$ . Is  $U_{32} = U_0$ ? I think there is a small error.
- 31 The *cosine of a matrix* is defined like  $e^A$ , by copying the series for  $\cos t$ :

$$\cos t = 1 - \frac{1}{2!}t^2 + \frac{1}{4!}t^4 - \dots \quad \cos A = I - \frac{1}{2!}A^2 + \frac{1}{4!}A^4 - \dots$$

- (a) If  $Ax = \lambda x$ , multiply each term times  $x$  to find the eigenvalue of  $\cos A$ .
- (b) Find the eigenvalues of  $A = \begin{bmatrix} \pi & \pi \\ \pi & \pi \end{bmatrix}$  with eigenvectors  $(1, 1)$  and  $(1, -1)$ . From the eigenvalues and eigenvectors of  $\cos A$ , find that matrix  $C = \cos A$ .
- (c) The second derivative of  $\cos(At)$  is  $-A^2 \cos(At)$ .

$$u(t) = \cos(At) u(0) \text{ solves } \frac{d^2 u}{dt^2} = -A^2 u \text{ starting from } u'(0) = 0.$$

Construct  $u(t) = \cos(At) u(0)$  by the usual three steps for that specific  $A$ :

1. Expand  $u(0) = (4, 2) = c_1 x_1 + c_2 x_2$  in the eigenvectors.
2. Multiply those eigenvectors by \_\_\_\_\_ and \_\_\_\_\_ (instead of  $e^{\lambda t}$ ).
3. Add up the solution  $u(t) = c_1 \text{_____} x_1 + c_2 \text{_____} x_2$ .

**32** Explain one of these three proofs that the square of  $e^A$  is  $e^{2A}$ .

1. Solving with  $e^A$  from  $t = 0$  to 1 and then 1 to 2 agrees with  $e^{2A}$  from 0 to 2.
2. The squared series  $(I + A + \frac{A^2}{2} + \cdots)^2$  matches  $I + 2A + \frac{(2A)^2}{2} + \cdots = e^{2A}$ .
3. If  $A$  can be diagonalized then  $(Xe^\Lambda X^{-1})(Xe^\Lambda X^{-1}) = Xe^{2\Lambda} X^{-1}$ .

### Notes on a Differential Equations Course

Certainly constant-coefficient linear equations are the simplest to solve. This section 6.3 of the book shows you part of a differential equations course, but there is more:

1. The second order equation  $mu'' + bu' + ku = 0$  has major importance in applications. The exponents  $\lambda$  in the solutions  $u = e^{\lambda t}$  solve  $m\lambda^2 + b\lambda + k = 0$ . The damping coefficient  $b$  is crucial:

**Underdamping**  $b^2 < 4mk$    **Critical damping**  $b^2 = 4mk$    **Overdamping**  $b^2 > 4mk$

This decides whether  $\lambda_1$  and  $\lambda_2$  are real roots or repeated roots or complex roots. With complex  $\lambda$ 's the solution  $u(t)$  oscillates as it decays.

2. Our equations had no forcing term  $f(t)$ . We were finding the “nullspace solution”. To  $u_n(t)$  we need to add a particular solution  $u_p(t)$  that balances the force  $f(t)$ :

$$\begin{array}{l} \text{Input } f(s) \text{ at time } s \\ \text{Growth factor } e^{A(t-s)} \\ \text{Add up outputs at time } t \end{array} \quad u_{\text{particular}} = \int_0^t e^{A(t-s)} f(s) ds.$$

This solution can also be discovered and studied by *Laplace transform*—that is the established way to convert linear differential equations to linear algebra.

In real applications, nonlinear differential equations are solved numerically. A standard method with good accuracy is “Runge-Kutta”—named after its discoverers. Analysis can find the constant solutions to  $du/dt = f(u)$ . Those are solutions  $u(t) = Y$  with  $f(Y) = 0$  and  $du/dt = 0$ : *no movement*. We can also understand stability or instability near  $u = Y$ . Far from  $Y$ , the computer takes over.

This basic course is the subject of my textbook (companion to this one) on *Differential Equations and Linear Algebra*: [math.mit.edu/dela](http://math.mit.edu/dela).

## 6.4 Symmetric Matrices

- 1 A symmetric matrix  $S$  has  $n$  **real eigenvalues**  $\lambda_i$  and  $n$  **orthonormal eigenvectors**  $q_1, \dots, q_n$ .
- 2 Every real symmetric  $S$  can be diagonalized:  $S = Q\Lambda Q^{-1} = Q\Lambda Q^T$
- 3 The number of positive eigenvalues of  $S$  equals the number of positive pivots.
- 4 Antisymmetric matrices  $A = -A^T$  have *imaginary*  $\lambda$ 's and *orthonormal (complex)*  $q$ 's.
- 5 Section 9.2 explains why the test  $S = S^T$  becomes  $S = \overline{S}^T$  for *complex matrices*.

$$S = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} = \overline{S}^T \text{ has real } \lambda = 1, -1. \quad A = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = -\overline{A}^T \text{ has } \lambda = i, -i.$$

It is no exaggeration to say that symmetric matrices  $S$  are the most important matrices the world will ever see—in the theory of linear algebra and also in the applications. We come immediately to the key question about symmetry. Not only the question, but also the two-part answer.

***What is special about  $Sx = \lambda x$  when  $S$  is symmetric?***

We look for special properties of the eigenvalues  $\lambda$  and eigenvectors  $x$  when  $S = S^T$ .

The diagonalization  $S = X\Lambda X^{-1}$  will reflect the symmetry of  $S$ . We get some hint by transposing to  $S^T = (X^{-1})^T \Lambda X^T$ . Those are the same since  $S = S^T$ . Possibly  $X^{-1}$  in the first form equals  $X^T$  in the second form? Then  $X^T X = I$ . That makes each eigenvector in  $X$  orthogonal to the other eigenvectors when  $S = S^T$ . Here are the key facts:

1. A symmetric matrix has only *real eigenvalues*.
2. The *eigenvectors* can be chosen *orthonormal*.

Those  $n$  orthonormal eigenvectors go into the columns of  $X$ . Every symmetric matrix can be diagonalized. ***Its eigenvector matrix  $X$  becomes an orthogonal matrix  $Q$ .*** Orthogonal matrices have  $Q^{-1} = Q^T$ —what we suspected about the eigenvector matrix is true. To remember it we write  $Q$  instead of  $X$ , when we choose orthonormal eigenvectors.

Why do we use the word “choose”? Because the eigenvectors do not *have* to be unit vectors. Their lengths are at our disposal. We will choose unit vectors—eigenvectors of length one, which are orthonormal and not just orthogonal. Then  $A = X\Lambda X^{-1}$  is in its special and particular form  $S = Q\Lambda Q^T$  for symmetric matrices.

**(Spectral Theorem)** Every symmetric matrix has the factorization  $S = Q\Lambda Q^T$  with real eigenvalues in  $\Lambda$  and orthonormal eigenvectors in the columns of  $Q$ :

**Symmetric diagonalization**  $S = Q\Lambda Q^{-1} = Q\Lambda Q^T \quad \text{with} \quad Q^{-1} = Q^T. \quad (1)$

It is easy to see that  $Q\Lambda Q^T$  is symmetric. Take its transpose. You get  $(Q^T)^T \Lambda^T Q^T$ , which is  $Q\Lambda Q^T$  again. The harder part is to prove that every symmetric matrix has real  $\lambda$ 's and orthonormal  $x$ 's. This is the “*spectral theorem*” in mathematics and the “*principal axis theorem*” in geometry and physics. We have to prove it! No choice. I will approach the proof in three steps:

1. By an example, showing real  $\lambda$ 's in  $\Lambda$  and orthonormal  $x$ 's in  $Q$ .
2. By a proof of those facts when no eigenvalues are repeated.
3. By a proof that allows repeated eigenvalues (at the end of this section).

**Example 1** Find the  $\lambda$ 's and  $x$ 's when  $S = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$  and  $S - \lambda I = \begin{bmatrix} 1 - \lambda & 2 \\ 2 & 4 - \lambda \end{bmatrix}$ .

**Solution** The determinant of  $S - \lambda I$  is  $\lambda^2 - 5\lambda$ . The eigenvalues are 0 and 5 (*both real*). We can see them directly:  $\lambda = 0$  is an eigenvalue because  $S$  is singular, and  $\lambda = 5$  matches the *trace* down the diagonal of  $S$ :  $0 + 5$  agrees with  $1 + 4$ .

Two eigenvectors are  $(2, -1)$  and  $(1, 2)$ —orthogonal but not yet orthonormal. The eigenvector for  $\lambda = 0$  is in the *nullspace* of  $A$ . The eigenvector for  $\lambda = 5$  is in the *column space*. We ask ourselves, why are the nullspace and column space perpendicular? The Fundamental Theorem says that the nullspace is perpendicular to the *row space*—not the column space. But our matrix is *symmetric*! Its row and column spaces are the same. Its eigenvectors  $(2, -1)$  and  $(1, 2)$  must be (and are) perpendicular.

These eigenvectors have length  $\sqrt{5}$ . Divide them by  $\sqrt{5}$  to get unit vectors. Put those unit eigenvectors into the columns of  $Q$ . Then  $Q^{-1}SQ$  is  $\Lambda$  and  $Q^{-1} = Q^T$ :

$$Q^{-1}SQ = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 5 \end{bmatrix} = \Lambda.$$

Now comes the  $n$  by  $n$  case. The  $\lambda$ 's are real when  $S = S^T$  and  $Sx = \lambda x$ .

**Real Eigenvalues** All the eigenvalues of a real symmetric matrix are real.

**Proof** Suppose that  $Sx = \lambda x$ . Until we know otherwise,  $\lambda$  might be a complex number  $a + ib$  ( $a$  and  $b$  real). Its complex conjugate is  $\bar{\lambda} = a - ib$ . Similarly the components of  $x$  may be complex numbers, and switching the signs of their imaginary parts gives  $\bar{x}$ .

The good thing is that  $\bar{\lambda}$  times  $\bar{x}$  is always the conjugate of  $\lambda$  times  $x$ . So we can take conjugates of  $Sx = \lambda x$ , remembering that  $S$  is real:

$$Sx = \lambda x \quad \text{leads to} \quad S\bar{x} = \bar{\lambda}\bar{x}. \quad \text{Transpose to} \quad \bar{x}^T S = \bar{x}^T \bar{\lambda}.$$

Now take the dot product of the first equation with  $\bar{x}$  and the last equation with  $x$ :

$$\bar{x}^T S x = \bar{x}^T \lambda x \quad \text{and also} \quad \bar{x}^T S x = \bar{x}^T \bar{\lambda} x. \quad (2)$$

The left sides are the same so the right sides are equal. One equation has  $\lambda$ , the other has  $\bar{\lambda}$ . They multiply  $\bar{x}^T x = |x_1|^2 + |x_2|^2 + \dots = \text{length squared}$  which is not zero. Therefore  $\lambda$  must equal  $\bar{\lambda}$ , and  $a + ib$  equals  $a - ib$ . So  $b = 0$  and  $\lambda = a = \text{real}$ . Q.E.D.

The eigenvectors come from solving the real equation  $(S - \lambda I)x = 0$ . So the  $x$ 's are also real. The important fact is that they are perpendicular.

**Orthogonal Eigenvectors** Eigenvectors of a real symmetric matrix (when they correspond to different  $\lambda$ 's) are always perpendicular.

**Proof** Suppose  $Sx = \lambda_1 x$  and  $Sy = \lambda_2 y$ . We are assuming here that  $\lambda_1 \neq \lambda_2$ . Take dot products of the first equation with  $y$  and the second with  $x$ :

$$\text{Use } S^T = S \quad (\lambda_1 x)^T y = (Sx)^T y = x^T S^T y = x^T S y = x^T \lambda_2 y. \quad (3)$$

The left side is  $x^T \lambda_1 y$ , the right side is  $x^T \lambda_2 y$ . Since  $\lambda_1 \neq \lambda_2$ , this proves that  $x^T y = 0$ . The eigenvector  $x$  (for  $\lambda_1$ ) is perpendicular to the eigenvector  $y$  (for  $\lambda_2$ ).

**Example 2** The eigenvectors of a 2 by 2 symmetric matrix have a special form:

$$\text{Not widely known} \quad S = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad \text{has} \quad x_1 = \begin{bmatrix} b \\ \lambda_1 - a \end{bmatrix} \quad \text{and} \quad x_2 = \begin{bmatrix} \lambda_2 - c \\ b \end{bmatrix}. \quad (4)$$

This is in the Problem Set. The point here is that  $x_1$  is perpendicular to  $x_2$ :

$$x_1^T x_2 = b(\lambda_2 - c) + (\lambda_1 - a)b = b(\lambda_1 + \lambda_2 - a - c) = 0.$$

This is zero because  $\lambda_1 + \lambda_2$  equals the trace  $a + c$ . Thus  $x_1^T x_2 = 0$ . Eagle eyes might notice the special case  $S = I$ , when  $b$  and  $\lambda_1 - a$  and  $\lambda_2 - c$  and  $x_1$  and  $x_2$  are all zero. Then  $\lambda_1 = \lambda_2 = 1$  is repeated. But of course  $S = I$  has perpendicular eigenvectors.

**Symmetric matrices  $S$  have orthogonal eigenvector matrices  $Q$ .** Look at this again:

$$\text{Symmetry} \quad S = X \Lambda X^{-1} \quad \text{becomes} \quad S = Q \Lambda Q^T \quad \text{with} \quad Q^T Q = I.$$

This says that every 2 by 2 symmetric matrix is (rotation)(stretch)(rotate back)

$$S = Q \Lambda Q^T = \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} q_1^T \\ q_2^T \end{bmatrix}. \quad (5)$$

Columns  $q_1$  and  $q_2$  multiply rows  $\lambda_1 q_1^T$  and  $\lambda_2 q_2^T$  to produce  $S = \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T$ .

**Every symmetric matrix**

$$S = Q\Lambda Q^T = \lambda_1 q_1 q_1^T + \cdots + \lambda_n q_n q_n^T. \quad (6)$$

Remember the steps to this great result (the spectral theorem).

*Section 6.2* Write  $Ax_i = \lambda_i x_i$  in matrix form  $AX = X\Lambda$  or  $A = X\Lambda X^{-1}$

*Section 6.4* Orthonormal  $x_i = q_i$  gives  $X = Q$   $S = Q\Lambda Q^{-1} = Q\Lambda Q^T$

$Q\Lambda Q^T$  in equation (6) has columns of  $Q\Lambda$  times rows of  $Q^T$ . Here is a direct proof.

**$S$  has correct eigenvectors  
Those  $q$ 's are orthonormal**

$$Sq_i = (\lambda_1 q_1 q_1^T + \cdots + \lambda_n q_n q_n^T) q_i = \lambda_i q_i. \quad (7)$$

## Complex Eigenvalues of Real Matrices

For any real matrix,  $Sx = \lambda x$  gives  $S\bar{x} = \bar{\lambda}\bar{x}$ . For a symmetric matrix,  $\lambda$  and  $x$  turn out to be real. Those two equations become the same. But a *nonsymmetric* matrix can easily produce  $\lambda$  and  $x$  that are complex. Then  $A\bar{x} = \bar{\lambda}\bar{x}$  is true but different from  $Ax = \lambda x$ . We get another complex eigenvalue (which is  $\bar{\lambda}$ ) and a new eigenvector (which is  $\bar{x}$ ):

*For real matrices, complex  $\lambda$ 's and  $x$ 's come in "conjugate pairs."*

$$\begin{array}{l} \lambda = a + ib \\ \bar{\lambda} = a - ib \end{array} \quad \text{If } Ax = \lambda x \text{ then } A\bar{x} = \bar{\lambda}\bar{x}. \quad (8)$$

**Example 3**  $A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$  has  $\lambda_1 = \cos \theta + i \sin \theta$  and  $\lambda_2 = \cos \theta - i \sin \theta$ .

Those eigenvalues are conjugate to each other. They are  $\lambda$  and  $\bar{\lambda}$ . The eigenvectors must be  $x$  and  $\bar{x}$ , because  $A$  is real:

$$\begin{array}{ll} \text{This is } \lambda x & Ax = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = (\cos \theta + i \sin \theta) \begin{bmatrix} 1 \\ -i \end{bmatrix} \\ \text{This is } \bar{\lambda} \bar{x} & A\bar{x} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = (\cos \theta - i \sin \theta) \begin{bmatrix} 1 \\ i \end{bmatrix}. \end{array} \quad (9)$$

Those eigenvectors  $(1, -i)$  and  $(1, i)$  are complex conjugates because  $A$  is real.

For this rotation matrix the absolute value is  $|\lambda| = 1$ , because  $\cos^2 \theta + \sin^2 \theta = 1$ . **This fact  $|\lambda| = 1$  holds for the eigenvalues of every orthogonal matrix  $Q$ .**

We apologize that a touch of complex numbers slipped in. They are unavoidable even when the matrix is real. Chapter 9 goes beyond complex numbers  $\lambda$  and complex eigenvectors  $x$  to complex matrices  $A$ . Then you have the whole picture.

We end with two optional discussions.

## Eigenvalues versus Pivots

The eigenvalues of  $A$  are very different from the pivots. For eigenvalues, we solve  $\det(A - \lambda I) = 0$ . For pivots, we use elimination. The only connection so far is this:

$$\text{product of pivots} = \text{determinant} = \text{product of eigenvalues.}$$

We are assuming a full set of pivots  $d_1, \dots, d_n$ . There are  $n$  real eigenvalues  $\lambda_1, \dots, \lambda_n$ . The  $d$ 's and  $\lambda$ 's are not the same, but they come from the same symmetric matrix. Those  $d$ 's and  $\lambda$ 's have a hidden relation. **For symmetric matrices the pivots and the eigenvalues have the same signs:**

*The number of positive eigenvalues of  $S = S^T$  equals the number of positive pivots.*

Special case:  $S$  has all  $\lambda_i > 0$  if and only if all pivots are positive.

That special case is an all-important fact for **positive definite matrices** in Section 6.5.

**Example 4** This symmetric matrix has one positive eigenvalue and one positive pivot:

$$\text{Matching signs} \quad S = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \quad \begin{array}{l} \text{has pivots 1 and } -8 \\ \text{eigenvalues 4 and } -2. \end{array}$$

The signs of the pivots match the signs of the eigenvalues, one plus and one minus. This could be false when the matrix is not symmetric:

$$\text{Opposite signs} \quad B = \begin{bmatrix} 1 & 6 \\ -1 & -4 \end{bmatrix} \quad \begin{array}{l} \text{has pivots 1 and 2} \\ \text{eigenvalues } -1 \text{ and } -2. \end{array}$$

The diagonal entries are a third set of numbers and we say nothing about them.

Here is a proof that the pivots and eigenvalues have matching signs, when  $S = S^T$ .

You see it best when the pivots are divided out of the rows of  $U$ . Then  $S$  is  $LDL^T$ . The diagonal pivot matrix  $D$  goes between triangular matrices  $L$  and  $L^T$ :

$$\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & -8 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \quad \text{This is } S = LDL^T. \text{ It is symmetric.}$$

*Watch the eigenvalues of  $LDL^T$  when  $L$  moves to  $I$ .  $S$  changes to  $D$ .*

The eigenvalues of  $LDL^T$  are 4 and  $-2$ . The eigenvalues of  $IDI^T$  are 1 and  $-8$  (the pivots!). The eigenvalues are changing, as the “3” in  $L$  moves to zero. But to change *sign*, a real eigenvalue would have to cross zero. The matrix would at that moment be singular. Our changing matrix always has pivots 1 and  $-8$ , so it is *never* singular. The signs cannot change, as the  $\lambda$ 's move to the  $d$ 's.

We repeat the proof for any  $S = LDL^T$ . Move  $L$  toward  $I$ , by moving the off-diagonal entries to zero. The pivots are not changing and not zero. The eigenvalues  $\lambda$  of  $LDL^T$  change to the eigenvalues  $d$  of  $IDI^T$ . Since these eigenvalues cannot cross zero as they move into the pivots, their signs cannot change. **Same signs for the  $\lambda$ 's and  $d$ 's.**

*This connects the two halves of applied linear algebra—pivots and eigenvalues.*

## All Symmetric Matrices are Diagonalizable

When no eigenvalues of  $A$  are repeated, the eigenvectors are sure to be independent. Then  $A$  can be diagonalized. But a repeated eigenvalue can produce a shortage of eigenvectors. This *sometimes* happens for nonsymmetric matrices. It *never* happens for symmetric matrices. **There are always enough eigenvectors to diagonalize  $S = S^T$ .**

Here is one idea for a proof. Change  $S$  slightly by a diagonal matrix  $\text{diag}(c, 2c, \dots, nc)$ . If  $c$  is very small, the new symmetric matrix will have no repeated eigenvalues. Then we know it has a full set of orthonormal eigenvectors. As  $c \rightarrow 0$  we obtain  $n$  orthonormal eigenvectors of the original  $S$ —even if some eigenvalues of that  $S$  are repeated.

Every mathematician knows that this argument is incomplete. How do we guarantee that the small diagonal matrix will separate the eigenvalues? (I am sure this is true.)

A different proof comes from a useful new factorization that applies to *all square matrices*  $A$ , symmetric or not. This new factorization quickly produces  $S = Q\Lambda Q^T$  with a full set of real orthonormal eigenvectors when  $S$  is any real symmetric matrix.

**Every square  $A$  factors into  $QTQ^{-1}$  where  $T$  is upper triangular and  $\bar{Q}^T = Q^{-1}$ .  
If  $A$  has real eigenvalues then  $Q$  and  $T$  can be chosen real:  $Q^T Q = I$ .**

This is *Schur's Theorem*. Its proof will go onto the website [math.mit.edu/linearalgebra](http://math.mit.edu/linearalgebra). Here I will show how  $T$  is diagonal ( $T = \Lambda$ ) when  $S$  is symmetric. Then  $S = Q\Lambda Q^T$ .

We know that every symmetric  $S$  has real eigenvalues, and Schur allows repeated  $\lambda$ 's:

Schur's  $S = QTQ^{-1}$  means that  $T = Q^T S Q$ . The transpose is again  $Q^T S Q$ .

The triangular  $T$  is symmetric when  $S^T = S$ . Then  $T$  must be diagonal and  $T = \Lambda$ .

This proves that  $S = Q\Lambda Q^{-1}$ . The symmetric  $S$  has  $n$  orthonormal eigenvectors in  $Q$ .

*Note.* I have added another proof in Section 7.2 of this book. That proof shows how the eigenvalues  $\lambda$  can be described *one at a time*. The largest  $\lambda_1$  is the maximum of  $x^T S x / x^T x$ . Then  $\lambda_2$  (second largest) is again the same maximum, if we only allow vectors  $x$  that are perpendicular to the first eigenvector. The third eigenvalue  $\lambda_3$  comes by requiring  $x^T q_1 = 0$  and  $x^T q_2 = 0 \dots$

This proof is placed in Chapter 7 because the same one-at-a-time idea succeeds for the *singular values of any matrix  $A$* . **Singular values come from  $A^T A$  and  $AA^T$ .**

### ■ REVIEW OF THE KEY IDEAS ■

1. Every symmetric matrix  $S$  has *real eigenvalues* and *perpendicular eigenvectors*.
2. Diagonalization becomes  $S = Q\Lambda Q^T$  with an orthogonal eigenvector matrix  $Q$ .
3. All symmetric matrices are diagonalizable, even with repeated eigenvalues.
4. The signs of the eigenvalues match the signs of the pivots, when  $S = S^T$ .
5. Every square matrix can be "triangularized" by  $A = QTQ^{-1}$ . If  $A = S$  then  $T = \Lambda$ .

### ■ WORKED EXAMPLES ■

**6.4 A** What matrix  $A$  has eigenvalues  $\lambda = 1, -1$  and eigenvectors  $\mathbf{x}_1 = (\cos \theta, \sin \theta)$  and  $\mathbf{x}_2 = (-\sin \theta, \cos \theta)$ ? Which of these properties can be predicted in advance?

$$A = A^T \quad A^2 = I \quad \det A = -1 \quad \text{pivot are + and -} \quad A^{-1} = A$$

**Solution** All those properties can be predicted! With real eigenvalues  $1, -1$  and orthonormal  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , the matrix  $A = Q\Lambda Q^T$  must be symmetric. The eigenvalues  $1$  and  $-1$  tell us that  $A^2 = I$  (since  $\lambda^2 = 1$ ) and  $A^{-1} = A$  (same thing) and  $\det A = -1$ . The two pivots must be positive and negative like the eigenvalues, since  $A$  is symmetric.

The matrix will be a reflection. Vectors in the direction of  $\mathbf{x}_1$  are unchanged by  $A$  (since  $\lambda = 1$ ). Vectors in the perpendicular direction are reversed (since  $\lambda = -1$ ). The reflection  $A = Q\Lambda Q^T$  is across the “ $\theta$ -line”. Write  $c$  for  $\cos \theta$  and  $s$  for  $\sin \theta$ :

$$A = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} c^2 - s^2 & 2cs \\ 2cs & s^2 - c^2 \end{bmatrix} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}.$$

Notice that  $\mathbf{x} = (1, 0)$  goes to  $A\mathbf{x} = (\cos 2\theta, \sin 2\theta)$  on the  $2\theta$ -line. And  $(\cos 2\theta, \sin 2\theta)$  goes back across the  $\theta$ -line to  $\mathbf{x} = (1, 0)$ .

**6.4 B** Find the eigenvalues and eigenvectors (discrete sines and cosines) of  $A_3$  and  $B_4$ .

$$A_3 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \quad B_4 = \begin{bmatrix} 1 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 1 \end{bmatrix}$$

The  $-1, 2, -1$  pattern in both matrices is a “second difference”. This is like a second derivative. Then  $A\mathbf{x} = \lambda\mathbf{x}$  and  $B\mathbf{x} = \lambda\mathbf{x}$  are like  $d^2x/dt^2 = \lambda x$ . This has eigenvectors  $x = \sin kt$  and  $x = \cos kt$  that are the bases for Fourier series.

$A_n$  and  $B_n$  lead to “discrete sines” and “discrete cosines” that are the bases for the *Discrete Fourier Transform*. This DFT is absolutely central to all areas of digital signal processing. The favorite choice for JPEG in image processing has been  $B_8$  of size  $n = 8$ .

**Solution** The eigenvalues of  $A_3$  are  $\lambda = 2 - \sqrt{2}$  and  $2$  and  $2 + \sqrt{2}$  (see 6.3 B). Their sum is  $6$  (the trace of  $A_3$ ) and their product is  $4$  (the determinant). The eigenvector matrix gives the “Discrete Sine Transform” and the eigenvectors fall onto sine curves.

$$\text{Sines} = \begin{bmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{bmatrix} \quad \text{Cosines} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \sqrt{2} - 1 & -1 & 1 - \sqrt{2} \\ 1 & 1 - \sqrt{2} & -1 & \sqrt{2} - 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

**Sine matrix = Eigenvectors of  $A_3$**

**Cosine matrix = Eigenvectors of  $B_4$**

The eigenvalues of  $B_4$  are  $\lambda = 2 - \sqrt{2}$  and  $2$  and  $2 + \sqrt{2}$  and  $0$  (the same as for  $A_3$ , plus the zero eigenvalue). The trace is still  $6$ , but the determinant is now zero. The eigenvector matrix  $C$  gives the 4-point “Discrete Cosine Transform”. The graph on the Web shows how the first two eigenvectors fall onto cosine curves. (So do all the eigenvectors of  $B$ .) These eigenvectors match cosines at the *halfway points*  $\pi/8, 3\pi/8, 5\pi/8, 7\pi/8$ .

## Problem Set 6.4

- 1 Which of these matrices  $ASB$  will be symmetric with eigenvalues 1 and  $-1$ ?

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & \\ & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$B = A^T$  doesn't do it.  $B = A^{-1}$  doesn't do it.  $B = \underline{\hspace{1cm}} = \underline{\hspace{1cm}}$  will succeed.  
So  $B$  must be an  $\underline{\hspace{1cm}}$  matrix.

- 2 Suppose  $S = S^T$ . When is  $ASB$  also symmetric with the same eigenvalues as  $S$ ?
- (a) Transpose  $ASB$  to see that it stays symmetric when  $B = \underline{\hspace{1cm}}$ .
- (b)  $ASB$  is similar to  $S$  (same eigenvalues) when  $B = \underline{\hspace{1cm}}$ .
- Put (a) and (b) together. The symmetric matrices similar to  $S$  look like  $(\underline{\hspace{1cm}})S(\underline{\hspace{1cm}})$ .
- 3 Write  $A$  as  $S + N$ , symmetric matrix  $S$  plus skew-symmetric matrix  $N$ :

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 4 & 3 & 0 \\ 8 & 6 & 5 \end{bmatrix} = S + N \quad (S^T = S \text{ and } N^T = -N).$$

For any square matrix,  $S = \frac{1}{2}(A + A^T)$  and  $N = \underline{\hspace{1cm}}$  add up to  $A$ .

- 4 If  $C$  is symmetric prove that  $A^T C A$  is also symmetric. (Transpose it.) When  $A$  is 6 by 3, what are the shapes of  $C$  and  $A^T C A$ ?
- 5 Find the eigenvalues and the unit eigenvectors of

$$S = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}.$$

- 6 Find an orthogonal matrix  $Q$  that diagonalizes  $S = \begin{bmatrix} -2 & 6 \\ 6 & 7 \end{bmatrix}$ . What is  $\Lambda$ ?
- 7 Find an orthogonal matrix  $Q$  that diagonalizes this symmetric matrix:

$$S = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -2 \\ 2 & -2 & 0 \end{bmatrix}.$$

- 8 Find *all* orthogonal matrices that diagonalize  $S = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$ .

- 9 (a) Find a symmetric matrix  $\begin{bmatrix} 1 & b \\ b & 1 \end{bmatrix}$  that has a negative eigenvalue.
- (b) How do you know it must have a negative pivot?
- (c) How do you know it can't have two negative eigenvalues?

- 10 If  $A^3 = 0$  then the eigenvalues of  $A$  must be \_\_\_\_\_. Give an example that has  $A \neq 0$ . But if  $A$  is symmetric, diagonalize it to prove that  $A$  must be a zero matrix.
- 11 If  $\lambda = a + ib$  is an eigenvalue of a real matrix  $A$ , then its conjugate  $\bar{\lambda} = a - ib$  is also an eigenvalue. (If  $Ax = \lambda x$  then also  $A\bar{x} = \bar{\lambda}\bar{x}$ : a conjugate pair  $\lambda$  and  $\bar{\lambda}$ .) Explain why every real 3 by 3 matrix has at least one real eigenvalue.
- 12 Here is a quick “proof” that the eigenvalues of every real matrix  $A$  are real:

$$\text{False proof } Ax = \lambda x \text{ gives } x^T Ax = \lambda x^T x \text{ so } \lambda = \frac{x^T Ax}{x^T x} = \frac{\text{real}}{\text{real}}.$$

Find the flaw in this reasoning—a hidden assumption that is not justified. You could test those steps on the  $90^\circ$  rotation matrix  $\begin{bmatrix} 0 & -1; & 1 & 0 \end{bmatrix}$  with  $\lambda = i$  and  $x = (i, 1)$ .

- 13 Write  $S$  and  $B$  in the form  $\lambda_1 x_1 x_1^T + \lambda_2 x_2 x_2^T$  of the spectral theorem  $Q\Lambda Q^T$ :

$$S = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix} \quad (\text{keep } \|x_1\| = \|x_2\| = 1).$$

- 14 Every 2 by 2 symmetric matrix is  $\lambda_1 x_1 x_1^T + \lambda_2 x_2 x_2^T = \lambda_1 P_1 + \lambda_2 P_2$ . Explain  $P_1 + P_2 = x_1 x_1^T + x_2 x_2^T = I$  from columns times rows of  $Q$ . Why is  $P_1 P_2 = 0$ ?
- 15 What are the eigenvalues of  $A = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$ ? Create a 4 by 4 antisymmetric matrix ( $A^T = -A$ ) and verify that all its eigenvalues are imaginary.
- 16 (Recommended) This matrix  $M$  is antisymmetric and also \_\_\_\_\_. Then all its eigenvalues are pure imaginary and they also have  $|\lambda| = 1$ . ( $\|Mx\| = \|x\|$  for every  $x$  so  $\|\lambda x\| = \|x\|$  for eigenvectors.) Find all four eigenvalues from the trace of  $M$ :

$$M = \frac{1}{\sqrt{3}} \begin{bmatrix} 0 & 1 & 1 & 1 \\ -1 & 0 & -1 & 1 \\ -1 & 1 & 0 & -1 \\ -1 & -1 & 1 & 0 \end{bmatrix} \quad \text{can only have eigenvalues } i \text{ or } -i.$$

- 17 Show that this  $A$  (**symmetric but complex**) has only one line of eigenvectors:

$$A = \begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix} \text{ is not even diagonalizable: eigenvalues } \lambda = 0, 0.$$

$A^T = A$  is not such a special property for complex matrices. The good property is  $\bar{A}^T = A$  (Section 9.2). Then all  $\lambda$ 's are real and the eigenvectors are orthogonal.

- 18 Even if  $A$  is rectangular, the block matrix  $S = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$  is symmetric:

$$Sx = \lambda x \text{ is } \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = \lambda \begin{bmatrix} y \\ z \end{bmatrix} \quad \text{which is } \begin{aligned} Az &= \lambda y \\ A^T y &= \lambda z. \end{aligned}$$

- (a) Show that  $-\lambda$  is also an eigenvalue, with the eigenvector  $(y, -z)$ .  
 (b) Show that  $A^T A z = \lambda^2 z$ , so that  $\lambda^2$  is an eigenvalue of  $A^T A$ .  
 (c) If  $A = I$  (2 by 2) find all four eigenvalues and eigenvectors of  $S$ .
- 19 If  $A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  in Problem 18, find all three eigenvalues and eigenvectors of  $S$ .
- 20 **Another proof that eigenvectors are perpendicular when  $S = S^T$ .** Two steps:
1. Suppose  $Sx = \lambda x$  and  $Sy = 0y$  and  $\lambda \neq 0$ . Then  $y$  is in the nullspace and  $x$  is in the column space. They are perpendicular because \_\_\_\_\_. Go carefully—why are these subspaces orthogonal?
  2. If  $Sy = \beta y$ , apply that argument to  $S - \beta I$ . One eigenvalue of  $S - \beta I$  moves to zero. The eigenvectors  $x, y$  stay the same—so they are perpendicular.
- 21 Find the eigenvector matrices  $Q$  for  $S$  and  $X$  for  $B$ . Show that  $X$  doesn't collapse at  $d = 1$ , even though  $\lambda = 1$  is repeated. Are those eigenvectors perpendicular?

$$S = \begin{bmatrix} 0 & d & 0 \\ d & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} -d & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & d \end{bmatrix} \quad \text{have } \lambda = 1, d, -d.$$

- 22 Write a 2 by 2 *complex* matrix with  $\bar{S}^T = S$  (a “Hermitian matrix”). Find  $\lambda_1$  and  $\lambda_2$  for your complex matrix. Check that  $\bar{x}_1^T x_2 = 0$  (this is complex orthogonality).
- 23 **True** (with reason) **or false** (with example).
- (a) A matrix with real eigenvalues and  $n$  real eigenvectors is symmetric.
  - (b) A matrix with real eigenvalues and  $n$  orthonormal eigenvectors is symmetric.
  - (c) The inverse of an invertible symmetric matrix is symmetric.
  - (d) The eigenvector matrix  $Q$  of a symmetric matrix is symmetric.
- 24 (A paradox for instructors) If  $AA^T = A^T A$  then  $A$  and  $A^T$  share the same eigenvectors (true).  $A$  and  $A^T$  always share the same eigenvalues. Find the flaw in this conclusion:  $A$  and  $A^T$  must have the same  $X$  and same  $\Lambda$ . Therefore  $A$  equals  $A^T$ .
- 25 (Recommended) Which of these classes of matrices do  $A$  and  $B$  belong to: Invertible, orthogonal, projection, permutation, diagonalizable, Markov?

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad B = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Which of these factorizations are possible for  $A$  and  $B$ :  $LU$ ,  $QR$ ,  $X\Lambda X^{-1}$ ,  $Q\Lambda Q^T$ ?

- 26 What number  $b$  in  $A = \begin{bmatrix} 2 & b \\ 1 & 0 \end{bmatrix}$  makes  $A = Q\Lambda Q^T$  possible? What number will make it impossible to diagonalize  $A$ ? What number makes  $A$  singular?

- 27 Find all 2 by 2 matrices that are orthogonal and also symmetric. Which two numbers can be eigenvalues of those two matrices?
- 28 This  $A$  is nearly symmetric. But its eigenvectors are far from orthogonal:

$$A = \begin{bmatrix} 1 & 10^{-15} \\ 0 & 1 + 10^{-15} \end{bmatrix} \quad \text{has eigenvectors} \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} ? \\ ? \end{bmatrix}$$

What is the angle between the eigenvectors?

- 29 (MATLAB) Take two symmetric matrices with different eigenvectors, say  $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$  and  $B = \begin{bmatrix} 8 & 1 \\ 1 & 0 \end{bmatrix}$ . Graph the eigenvalues  $\lambda_1(A + tB)$  and  $\lambda_2(A + tB)$  for  $-8 < t < 8$ . Peter Lax says on page 113 of *Linear Algebra* that  $\lambda_1$  and  $\lambda_2$  appear to be on a collision course at certain values of  $t$ . “Yet at the last minute they turn aside.” How close does  $\lambda_1$  come to  $\lambda_2$ ?

### Challenge Problems

- 30 For complex matrices, the symmetry  $S^T = S$  that produces real eigenvalues must change in Section 9.2 to  $\overline{S}^T = S$ . From  $\det(S - \lambda I) = 0$ , find the eigenvalues of the 2 by 2 **Hermitian matrix**  $S = \begin{bmatrix} 4 & 2 + i \\ 2 - i & 0 \end{bmatrix} = \overline{S}^T$ .
- 31 **Normal matrices** have  $\overline{N}^T N = N \overline{N}^T$ . For real matrices, this is  $N^T N = N N^T$ . Normal includes symmetric, skew-symmetric, and orthogonal (with real  $\lambda$ , imaginary  $\lambda$ , and  $|\lambda| = 1$ ). Other normal matrices can have any complex eigenvalues.

Key point: **Normal matrices have  $n$  orthonormal eigenvectors.** Those vectors  $x_i$  probably will have complex components. In that complex case (Chapter 9) orthogonality means  $\overline{x_i}^T x_j = 0$ . Inner products (dot products)  $x^T y$  become  $\overline{x}^T y$ .

**The test for  $n$  orthonormal columns in  $Q$  becomes  $\overline{Q}^T Q = I$  instead of  $Q^T Q = I$ .**

$N$  has  $n$  **orthonormal eigenvectors** ( $N = Q \Lambda \overline{Q}^T$ ) if and only if  $N$  is **normal**.

- (a) Start from  $N = Q \Lambda \overline{Q}^T$  with  $\overline{Q}^T Q = I$ . Show that  $\overline{N}^T N = N \overline{N}^T$ :  $N$  is normal.
- (b) Now start from  $\overline{N}^T N = N \overline{N}^T$ . Schur found  $A = Q T \overline{Q}^T$  for every matrix  $A$ , with a *triangular*  $T$ . For normal matrices  $A = N$  we must show (in 3 steps) that this triangular matrix  $T$  will actually be diagonal. Then  $T = \Lambda$ .

Step 1. Put  $N = Q T \overline{Q}^T$  into  $\overline{N}^T N = N \overline{N}^T$  to find  $\overline{T}^T T = T \overline{T}^T$ .

Step 2. Suppose  $T = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$  has  $\overline{T}^T T = T \overline{T}^T$ . Prove that  $b = 0$ .

Step 3. Extend Step 2 to size  $n$ . Any normal triangular  $T$  must be diagonal.

**32** If  $\lambda_{\max}$  is the largest eigenvalue of a symmetric matrix  $S$ , no diagonal entry can be larger than  $\lambda_{\max}$ . What is the first entry  $a_{11}$  of  $S = Q\Lambda Q^T$ ? Show why  $a_{11} \leq \lambda_{\max}$ .

**33** Suppose  $A^T = -A$  (real *antisymmetric* matrix). Explain these facts about  $A$ :

- (a)  $x^T Ax = 0$  for every real vector  $x$ .
- (b) The eigenvalues of  $A$  are pure imaginary.
- (c) The determinant of  $A$  is positive or zero (not negative).

For (a), multiply out an example of  $x^T Ax$  and watch terms cancel. Or reverse  $x^T(Ax)$  to  $-(Ax)^T x$ . For (b),  $Az = \lambda z$  leads to  $\bar{z}^T Az = \lambda \bar{z}^T z = \lambda \|z\|^2$ . Part (a) shows that  $\bar{z}^T Az = (x - iy)^T A(x + iy)$  has zero real part. Then (b) helps with (c).

**34** If  $S$  is symmetric and all its eigenvalues are  $\lambda = 2$ , how do you know that  $S$  must be  $2I$ ? Key point: Symmetry guarantees that  $S = Q\Lambda Q^T$ . What is that  $\Lambda$ ?

**35** Which symmetric matrices  $S$  are also orthogonal? Then  $S^T = S^{-1}$ .

- (a) Show how symmetry and orthogonality lead to  $S^2 = I$ .
- (b) What are the possible eigenvalues of this  $S$ ?
- (c) What are the possible eigenvalue matrices  $\Lambda$ ? Then  $S$  must be  $Q\Lambda Q^T$  for those  $\Lambda$  and any orthogonal  $Q$ .

**36** If  $S$  is symmetric, show that  $A^T SA$  is also symmetric (take the transpose of  $A^T SA$ ). Here  $A$  is  $m$  by  $n$  and  $S$  is  $m$  by  $m$ . Are eigenvalues of  $S$  = eigenvalues of  $A^T SA$ ?

In case  $A$  is square and invertible,  $A^T SA$  is called *congruent* to  $S$ . They have the same number of positive, negative, and zero eigenvalues: **Law of Inertia**.

**37** Here is a way to show that  $a$  is *in between* the eigenvalues  $\lambda_1$  and  $\lambda_2$  of  $S$ :

$$S = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad \det(S - \lambda I) = \lambda^2 - a\lambda - c\lambda + ac - b^2$$

is a parabola opening upwards (because of  $\lambda^2$ )

Show that  $\det(S - \lambda I)$  is negative at  $\lambda = a$ . So the parabola crosses the axis left and right of  $\lambda = a$ . It crosses at the two eigenvalues of  $S$  so they must enclose  $a$ .

The  $n - 1$  eigenvalues of  $A$  always fall between the  $n$  eigenvalues of  $S = \begin{bmatrix} A & b \\ b^T & c \end{bmatrix}$ .

## 6.5 Positive Definite Matrices

- 1 **Symmetric  $S$**  : all eigenvalues  $> 0 \Leftrightarrow$  all pivots  $> 0 \Leftrightarrow$  all upper left determinants  $> 0$ .
- 2 The matrix  $S$  is then **positive definite**. The energy test is  $\mathbf{x}^T S \mathbf{x} > 0$  for all vectors  $\mathbf{x} \neq \mathbf{0}$ .
- 3 One more test for positive definiteness:  $S = A^T A$  with independent columns in  $A$ .
- 4 **Positive semidefinite  $S$**  allows  $\lambda = 0$ , pivot = 0, determinant = 0, energy  $\mathbf{x}^T S \mathbf{x} = 0$ .
- 5 The equation  $\mathbf{x}^T S \mathbf{x} = 1$  gives an ellipse in  $\mathbf{R}^n$  when  $S$  is symmetric positive definite.

This section concentrates on *symmetric matrices that have positive eigenvalues*. If symmetry makes a matrix important, this extra property (*all  $\lambda > 0$* ) makes it truly special. When we say special, we don't mean rare. Symmetric matrices with positive eigenvalues are at the center of all kinds of applications. They are called **positive definite**.

The first problem is to recognize positive definite matrices. You may say, just find all the eigenvalues and test  $\lambda > 0$ . That is exactly what we want to avoid. Calculating eigenvalues is work. When the  $\lambda$ 's are needed, we can compute them. But if we just want to know that all the  $\lambda$ 's are positive, there are faster ways. Here are two goals of this section:

- To find *quick tests* on a symmetric matrix that guarantee *positive eigenvalues*.
- To explain important applications of positive definiteness.

Every eigenvalue is real because the matrix is symmetric.

*Start with 2 by 2. When does  $S = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$  have  $\lambda_1 > 0$  and  $\lambda_2 > 0$ ?*

**Test:** The eigenvalues of  $S$  are positive if and only if  $a > 0$  and  $ac - b^2 > 0$ .

$S_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$  is **not** positive definite because  $ac - b^2 = 1 - 4 < 0$

$S_2 = \begin{bmatrix} 1 & -2 \\ -2 & 6 \end{bmatrix}$  is positive definite because  $a = 1$  and  $ac - b^2 = 6 - 4 > 0$

$S_3 = \begin{bmatrix} -1 & 2 \\ 2 & -6 \end{bmatrix}$  is **not** positive definite (even with  $\det A = +2$ ) because  $a = -1$

The eigenvalues 3 and  $-1$  of  $S_1$  confirm that  $S_1$  is *not* positive definite. Positive trace  $3 - 1 = 2$ , but negative determinant  $(3)(-1) = -3$ . And  $S_3 = -S_2$  is *negative* definite. Two positive eigenvalues for  $S_2$ , two negative eigenvalues for  $S_3$ .

*Proof that the 2 by 2 test is passed when  $\lambda_1 > 0$  and  $\lambda_2 > 0$ .* Their product  $\lambda_1 \lambda_2$  is the determinant so  $ac - b^2 > 0$ . Their sum  $\lambda_1 + \lambda_2$  is the trace so  $a + c > 0$ . Then  $a$  and  $c$  are

both positive (if  $a$  or  $c$  is not positive,  $ac - b^2 > 0$  will fail). Problem 1 reverses the reasoning to show that the tests  $a > 0$  and  $ac > b^2$  guarantee  $\lambda_1 > 0$  and  $\lambda_2 > 0$ .

This test uses the 1 by 1 determinant  $a$  and the 2 by 2 determinant  $ac - b^2$ . When  $S$  is 3 by 3,  $\det S > 0$  is the third part of the test. The next test requires *positive pivots*.

**Test: The eigenvalues of  $S$  are positive if and only if the pivots are positive:**

$$a > 0 \quad \text{and} \quad \frac{ac - b^2}{a} > 0.$$

$a > 0$  is required in both tests. So  $ac > b^2$  is also required, for the determinant test and now the pivot test. The point is to recognize that ratio as the *second pivot* of  $S$ :

$$\begin{array}{ccc} \begin{bmatrix} a & b \\ b & c \end{bmatrix} & \begin{array}{l} \text{The first pivot is } a \\ \longrightarrow \\ \text{The multiplier is } b/a \end{array} & \begin{bmatrix} a & b \\ 0 & c - \frac{b}{a}b \end{bmatrix} \end{array} \quad \begin{array}{l} \text{The second pivot is} \\ c - \frac{b^2}{a} = \frac{ac - b^2}{a} \end{array}$$

This connects two big parts of linear algebra. **Positive eigenvalues mean positive pivots and vice versa.** Each pivot is a ratio of upper left determinants. The pivots give a quick test for  $\lambda > 0$ , and they are a lot faster to compute than the eigenvalues. It is very satisfying to see pivots and determinants and eigenvalues come together in this course.

$$\text{3 by 3 example} \quad S = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \text{ is positive definite} \quad \begin{array}{l} \text{eigenvalues } 1, 1, 4 \\ \text{determinants } 2 \text{ and } 3 \text{ and } 4 \\ \text{pivots } 2 \text{ and } 3/2 \text{ and } 4/3 \end{array}$$

$S - I$  will be *semidefinite*: eigenvalues 0, 0, 3.  $S - 2I$  is *indefinite* because  $\lambda = -1, -1, 2$ .

Now comes a different way to look at symmetric matrices with positive eigenvalues.

### Energy-based Definition

From  $Sx = \lambda x$ , multiply by  $x^T$  to get  $x^T Sx = \lambda x^T x$ . The right side is a positive  $\lambda$  times a positive number  $x^T x = \|x\|^2$ . So the left side  $x^T Sx$  is positive for any eigenvector.

**Important point:** The new idea is that  $x^T Sx$  is *positive for all nonzero vectors  $x$* , not just the eigenvectors. In many applications this number  $x^T Sx$  (or  $\frac{1}{2}x^T Sx$ ) is the **energy** in the system. The requirement of positive energy gives *another definition* of a positive definite matrix. I think this energy-based definition is the fundamental one.

Eigenvalues and pivots are two equivalent ways to test the new requirement  $x^T Sx > 0$ .

**Definition**  $S$  is positive definite if  $x^T S x > 0$  for every nonzero vector  $x$ :

$$\text{2 by 2} \quad x^T S x = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2 > 0. \quad (1)$$

The four entries  $a, b, b, c$  give the four parts of  $x^T S x$ . From  $a$  and  $c$  come the pure squares  $ax^2$  and  $cy^2$ . From  $b$  and  $b$  off the diagonal come the cross terms  $bxy$  and  $byx$  (the same). Adding those four parts gives  $x^T S x$ . This energy-based definition leads to a basic fact:

*If  $S$  and  $T$  are symmetric positive definite, so is  $S + T$ .*

**Reason:**  $x^T(S+T)x$  is simply  $x^T S x + x^T T x$ . Those two terms are positive (for  $x \neq 0$ ) so  $S + T$  is also positive definite. The pivots and eigenvalues are not easy to follow when matrices are added, but the energies just add.

$x^T S x$  also connects with our final way to recognize a positive definite matrix. Start with any matrix  $A$ , possibly rectangular. We know that  $S = A^T A$  is square and symmetric. More than that,  $S$  will be positive definite when  $A$  has independent columns:

*Test: If the columns of  $A$  are independent, then  $S = A^T A$  is positive definite.*

Again eigenvalues and pivots are not easy. But the number  $x^T S x$  is the same as  $x^T A^T A x$ .  $x^T A^T A x$  is exactly  $(Ax)^T(Ax) = \|Ax\|^2$ —another important proof by parenthesis! That vector  $Ax$  is not zero when  $x \neq 0$  (this is the meaning of independent columns). Then  $x^T S x$  is the positive number  $\|Ax\|^2$  and the matrix  $S$  is positive definite.

Let me collect this theory together, into *five equivalent statements* of positive definiteness. You will see how that key idea connects the whole subject of linear algebra: pivots, determinants, eigenvalues, and least squares (from  $A^T A$ ). Then come the applications.

*When a symmetric matrix  $S$  has one of these five properties, it has them all :*

1. All  $n$  pivots of  $S$  are positive.
2. All  $n$  upper left determinants are positive.
3. All  $n$  eigenvalues of  $S$  are positive.
4.  $x^T S x$  is positive except at  $x = 0$ . This is the *energy-based* definition.
5.  $S$  equals  $A^T A$  for a matrix  $A$  with *independent columns*.

The “upper left determinants” are 1 by 1, 2 by 2,  $\dots$ ,  $n$  by  $n$ . The last one is the determinant of the complete matrix  $S$ . This theorem ties together the whole linear algebra course.

**Example 1** Test these symmetric matrices  $S$  and  $T$  for positive definiteness:

$$S = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} 2 & -1 & b \\ -1 & 2 & -1 \\ b & -1 & 2 \end{bmatrix}.$$

**Solution** The pivots of  $S$  are 2 and  $\frac{3}{2}$  and  $\frac{4}{3}$ , all positive. Its upper left determinants are 2 and 3 and 4, all positive. The eigenvalues of  $S$  are  $2 - \sqrt{2}$  and  $2 + \sqrt{2}$ , all positive. That completes tests 1, 2, and 3. Any one test is decisive!

I have three candidates  $A_1, A_2, A_3$  to suggest for  $S = A^T A$ . They all show that  $S$  is positive definite.  $A_1$  is a first difference matrix, 4 by 3, to produce  $-1, 2, -1$  in  $S$ :

$$S = A_1^T A_1 \quad \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

The three columns of  $A_1$  are independent. Therefore  $S$  is positive definite.

$A_2$  comes from  $S = LDL^T$  (the symmetric version of  $S = LU$ ). Elimination gives the pivots  $2, \frac{3}{2}, \frac{4}{3}$  in  $D$  and the multipliers  $-\frac{1}{2}, 0, -\frac{2}{3}$  in  $L$ . **Just put  $A_2 = L\sqrt{D}$ .**

$$LDL^T = \begin{bmatrix} 1 & & \\ -\frac{1}{2} & 1 & \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & \frac{3}{2} & \\ & & \frac{4}{3} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ & 1 & -\frac{2}{3} \\ & & 1 \end{bmatrix} = (L\sqrt{D})(L\sqrt{D})^T = A_2^T A_2. \quad \text{A}_2 \text{ is the Cholesky factor of } S$$

This triangular choice of  $A$  has square roots (not so beautiful). It is the “Cholesky factor” of  $S$  and the MATLAB command is  $A = \text{chol}(S)$ . In applications, the rectangular  $A_1$  is how we build  $S$  and this Cholesky  $A_2$  is how we break it apart.

**Eigenvalues give the symmetric choice  $A_3 = Q\sqrt{\Lambda}Q^T$ .** This is also successful with  $A_3^T A_3 = Q\Lambda Q^T = S$ . All tests show that the  $-1, 2, -1$  matrix  $S$  is positive definite.

To see that the energy  $x^T S x$  is positive, we can write it as a sum of squares. The three choices  $A_1, A_2, A_3$  give three different ways to split up  $x^T S x$ :

$$x^T S x = 2x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_2x_3 + 2x_3^2 \quad \text{Rewrite with squares}$$

$$||A_1 x||^2 = x_1^2 + (x_2 - x_1)^2 + (x_3 - x_2)^2 + x_3^2 \quad \text{Using differences in } A_1$$

$$||A_2 x||^2 = 2(x_1 - \frac{1}{2}x_2)^2 + \frac{3}{2}(x_2 - \frac{2}{3}x_3)^2 + \frac{4}{3}x_3^2 \quad \text{Using } S = LDL^T$$

$$||A_3 x||^2 = \lambda_1(q_1^T x)^2 + \lambda_2(q_2^T x)^2 + \lambda_3(q_3^T x)^2 \quad \text{Using } S = Q\Lambda Q^T$$

Now turn to  $T$  (top of this page). The  $(1, 3)$  and  $(3, 1)$  entries move away from 0 to  $b$ . This  $b$  must not be too large! *The determinant test is easiest.* The 1 by 1 determinant is 2, the 2 by 2 determinant  $T$  is still 3. The 3 by 3 determinant involves  $b$ :

$$\text{Test on } T \quad \det T = 4 + 2b - 2b^2 = (1 + b)(4 - 2b) \quad \text{must be positive.}$$

At  $b = -1$  and  $b = 2$  we get  $\det T = 0$ . *Between  $b = -1$  and  $b = 2$  this matrix  $T$  is positive definite.* The corner entry  $b = 0$  in the matrix  $S$  was safely between  $-1$  and  $2$ .

### Positive Semidefinite Matrices

Often we are at the edge of positive definiteness. The determinant is zero. The smallest eigenvalue is zero. The energy in its eigenvector is  $\mathbf{x}^T S \mathbf{x} = \mathbf{x}^T \mathbf{0} \mathbf{x} = 0$ . These matrices on the edge are called *positive semidefinite*. Here are two examples (not invertible):

$$S = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \text{ and } T = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \text{ are positive semidefinite.}$$

$S$  has eigenvalues 5 and 0. Its upper left determinants are 1 and 0. Its rank is only 1. This matrix  $S$  factors into  $A^T A$  with **dependent columns** in  $A$ :

$$\begin{array}{ll} \text{Dependent columns in } A & \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = A^T A. \\ \text{Positive semidefinite } S & \end{array}$$

If 4 is increased by any small number, the matrix  $S$  will become positive definite.

The cyclic  $T$  also has zero determinant (computed above when  $b = -1$ ).  $T$  is singular. The eigenvector  $\mathbf{x} = (1, 1, 1)$  has  $T\mathbf{x} = \mathbf{0}$  and energy  $\mathbf{x}^T T \mathbf{x} = 0$ . Vectors  $\mathbf{x}$  in all other directions do give positive energy. This  $T$  can be written as  $A^T A$  in many ways, but  $A$  will always have *dependent* columns, with  $(1, 1, 1)$  in its nullspace:

$$\begin{array}{ll} \text{Second differences } T & \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}, \\ \text{from first differences } A & \\ \text{Cyclic } T \text{ from cyclic } A & \end{array}$$

Positive semidefinite matrices have all  $\lambda \geq 0$  and all  $\mathbf{x}^T S \mathbf{x} \geq 0$ . Those weak inequalities ( $\geq$  instead of  $>$ ) include positive definite  $S$  and also the singular matrices at the edge.

### The Ellipse $ax^2 + 2bxy + cy^2 = 1$

Think of a tilted ellipse  $\mathbf{x}^T S \mathbf{x} = 1$ . Its center is  $(0, 0)$ , as in Figure 6.7a. Turn it to line up with the coordinate axes ( $X$  and  $Y$  axes). That is Figure 6.7b. These two pictures show the geometry behind the factorization  $S = Q\Lambda Q^{-1} = Q\Lambda Q^T$ :

1. The tilted ellipse is associated with  $S$ . Its equation is  $\mathbf{x}^T S \mathbf{x} = 1$ .
2. The lined-up ellipse is associated with  $\Lambda$ . Its equation is  $\mathbf{X}^T \Lambda \mathbf{X} = 1$ .
3. The rotation matrix that lines up the ellipse is the eigenvector matrix  $Q$ .

**Example 2** Find the axes of this tilted ellipse  $5x^2 + 8xy + 5y^2 = 1$ .

**Solution** Start with the positive definite matrix that matches this equation:

$$\text{The equation is } \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1. \quad \text{The matrix is } S = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}.$$

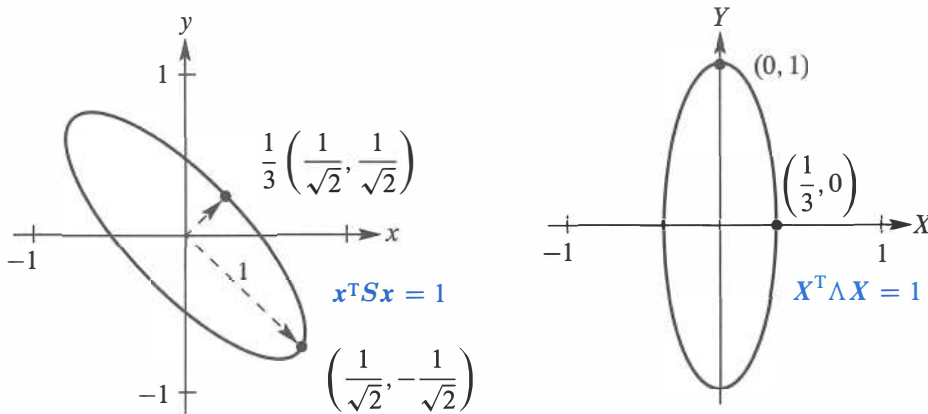


Figure 6.7: The tilted ellipse  $5x^2 + 8xy + 5y^2 = 1$ . Lined up it is  $9X^2 + Y^2 = 1$ .

The eigenvectors are  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  and  $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ . Divide by  $\sqrt{2}$  for unit vectors. Then  $S = Q\Lambda Q^T$ :

$$\begin{array}{l} \text{Eigenvectors in } Q \\ \text{Eigenvalues 9 and 1} \end{array} \quad \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Now multiply by  $\begin{bmatrix} x & y \end{bmatrix}$  on the left and  $\begin{bmatrix} x \\ y \end{bmatrix}$  on the right to get  $x^T S x = (x^T Q) \Lambda (Q^T x)$ :

$$x^T S x = \text{sum of squares} \quad 5x^2 + 8xy + 5y^2 = 9 \left( \frac{x+y}{\sqrt{2}} \right)^2 + 1 \left( \frac{x-y}{\sqrt{2}} \right)^2. \quad (2)$$

The coefficients are not the pivots 5 and 9/5 from  $D$ , they are the eigenvalues 9 and 1 from  $\Lambda$ . Inside the squares are the eigenvectors  $q_1 = (1, 1)/\sqrt{2}$  and  $q_2 = (1, -1)/\sqrt{2}$ .

**The axes of the tilted ellipse point along those eigenvectors.** This explains why  $S = Q\Lambda Q^T$  is called the “principal axis theorem”—it displays the axes. Not only the axis directions (from the eigenvectors) but also the axis lengths (from the eigenvalues). To see it all, use capital letters for the new coordinates that line up the ellipse:

$$\text{Lined up} \quad \frac{x+y}{\sqrt{2}} = X \quad \text{and} \quad \frac{x-y}{\sqrt{2}} = Y \quad \text{and} \quad 9X^2 + Y^2 = 1.$$

The largest value of  $X^2$  is  $1/9$ . The endpoint of the shorter axis has  $X = 1/3$  and  $Y = 0$ . Notice: The *bigger* eigenvalue  $\lambda_1$  gives the *shorter* axis, of half-length  $1/\sqrt{\lambda_1} = 1/3$ . The smaller eigenvalue  $\lambda_2 = 1$  gives the greater length  $1/\sqrt{\lambda_2} = 1$ .

In the  $xy$  system, the axes are along the eigenvectors of  $S$ . In the  $XY$  system, the axes are along the eigenvectors of  $\Lambda$ —the coordinate axes. All comes from  $S = Q\Lambda Q^T$ .

$S = Q\Lambda Q^T$  is positive definite when all  $\lambda_i > 0$ . The graph of  $x^T S x = 1$  is an ellipse:

$$\text{Ellipse } \begin{bmatrix} x & y \end{bmatrix} Q\Lambda Q^T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X & Y \end{bmatrix} \Lambda \begin{bmatrix} X \\ Y \end{bmatrix} = \lambda_1 X^2 + \lambda_2 Y^2 = 1. \quad (3)$$

The axes point along eigenvectors of  $S$ . The half-lengths are  $1/\sqrt{\lambda_1}$  and  $1/\sqrt{\lambda_2}$ .

$S = I$  gives the circle  $x^2 + y^2 = 1$ . If one eigenvalue is negative (exchange 4's and 5's in  $S$ ), the ellipse changes to a *hyperbola*. The sum of squares becomes a *difference of squares*:  $9X^2 - Y^2 = 1$ . For a negative definite matrix like  $S = -I$ , with both  $\lambda$ 's negative, the graph of  $-x^2 - y^2 = 1$  has no points at all.

If  $S$  is  $n$  by  $n$ ,  $x^T S x = 1$  is an “ellipsoid” in  $\mathbf{R}^n$ . Its axes are the eigenvectors of  $S$ .

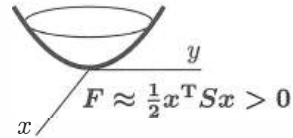
### Important Application: Test for a Minimum

Does  $F(x, y)$  have a minimum if  $\partial F/\partial x = 0$  and  $\partial F/\partial y = 0$  at the point  $(x, y) = (0, 0)$ ?

For  $f(x)$ , the test for a minimum comes from calculus:  $df/dx$  is zero and  $d^2 f/dx^2 > 0$ . Two variables in  $F(x, y)$  produce a symmetric matrix  $S$ . It contains *four second derivatives*. **Positive  $d^2 f/dx^2$  changes to positive definite  $S$ :**

**Second  
derivatives**

$$S = \begin{bmatrix} \partial^2 F/\partial x^2 & \partial^2 F/\partial x \partial y \\ \partial^2 F/\partial y \partial x & \partial^2 F/\partial y^2 \end{bmatrix}$$



$F(x, y)$  has a minimum if  $\partial F/\partial x = \partial F/\partial y = 0$  and  $S$  is positive definite.

Reason:  $S$  reveals the all-important terms  $ax^2 + 2bxy + cy^2$  near  $(x, y) = (0, 0)$ . The second derivatives of  $F$  are  $2a, 2b, 2b, 2c$ . For  $F(x, y, z)$  the matrix  $S$  will be 3 by 3.

### ■ REVIEW OF THE KEY IDEAS ■

1. Positive definite matrices have positive eigenvalues and positive pivots.
2. A quick test is given by the upper left determinants:  $a > 0$  and  $ac - b^2 > 0$ .
3. The graph of the energy  $x^T S x$  is then a “bowl” going up from  $x = 0$ :  

$$x^T S x = ax^2 + 2bxy + cy^2$$
 is positive except at  $(x, y) = (0, 0)$ .
4.  $S = A^T A$  is automatically positive definite if  $A$  has independent columns.
5. The ellipsoid  $x^T S x = 1$  has its axes along the eigenvectors of  $S$ . Lengths  $1/\sqrt{\lambda}$ .
6. Minimum of  $F(x, y)$  if  $\frac{\partial F}{\partial x} = \frac{\partial F}{\partial y} = 0$  and 2nd derivative matrix is positive definite.

### ■ WORKED EXAMPLES ■

**6.5 A** The great factorizations of a symmetric matrix are  $S = LDL^T$  from pivots and multipliers, and  $S = Q\Lambda Q^T$  from eigenvalues and eigenvectors. Try these  $n$  by  $n$  tests on `pascal(6)` and `ones(6)` and `hilb(6)` and other matrices in MATLAB's gallery.

**pascal(6)** is positive *definite* because all its pivots are 1 (Worked Example 2.6 A).

**ones(6)** is positive *semidefinite* because its eigenvalues are 0, 0, 0, 0, 0, 6.

**H=hilb(6)** is positive *definite* even though `eig(H)` shows eigenvalues very near zero.

**Hilbert matrix**  $x^T H x = \int_0^1 (x_1 + x_2 s + \cdots + x_6 s^5)^2 ds > 0$ ,  $H_{ij} = 1/(i + j - 1)$ .

**rand(6) + rand(6)'** can be positive definite or not. *Experiments gave only 2 in 20000.*

$n = 20000$ ;  $p = 0$ ; for  $k = 1:n$ ,  $A = \text{rand}(6)$ ;  $p = p + \text{all}(\text{eig}(A + A') > 0)$ ; end,  $p/n$

**6.5 B** When is the symmetric block matrix  $M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$  positive definite?

**Solution** Multiply the first row of  $M$  by  $B^T A^{-1}$  and subtract from the second row, to get a block of zeros. The *Schur complement*  $S = C - B^T A^{-1} B$  appears in the corner:

$$\begin{bmatrix} I & 0 \\ -B^T A^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & C - B^T A^{-1} B \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & S \end{bmatrix} \quad (4)$$

Those two blocks  $A$  and  $S$  must be positive definite. Their pivots are the pivots of  $M$ .

**6.5 C** Find the eigenvalues of the  $-1, 2, -1$  tridiagonal  $n$  by  $n$  matrix  $S$  (my favorite).

**Solution** The best way is to guess  $\lambda$  and  $x$ . Then check  $Sx = \lambda x$ . Guessing could not work for most matrices, but special cases are a big part of mathematics (pure and applied).

The key is hidden in a differential equation. The second difference matrix  $S$  is like a *second derivative*, and those eigenvalues are much easier to see:

|                                           |                                     |      |            |     |
|-------------------------------------------|-------------------------------------|------|------------|-----|
| Eigenvalues $\lambda_1, \lambda_2, \dots$ | $\frac{d^2 y}{dx^2} = \lambda y(x)$ | with | $y(0) = 0$ | (5) |
| Eigenfunctions $y_1, y_2, \dots$          |                                     |      | $y(1) = 0$ |     |

Try  $y = \sin cx$ . Its second derivative is  $y'' = -c^2 \sin cx$ . So the eigenvalue in (5) will be  $\lambda = -c^2$ , provided  $y = \sin cx$  satisfies the end point conditions  $y(0) = 0 = y(1)$ .

Certainly  $\sin 0 = 0$  (this is where cosines are eliminated). At the other end  $x = 1$ , we need  $y(1) = \sin c = 0$ . The number  $c$  must be  $k\pi$ , a multiple of  $\pi$ . Then  $\lambda$  is  $-k^2\pi^2$ :

|                                   |                                                          |     |
|-----------------------------------|----------------------------------------------------------|-----|
| Eigenvalues $\lambda = -k^2\pi^2$ | $\frac{d^2}{dx^2} \sin k\pi x = -k^2\pi^2 \sin k\pi x$ . | (6) |
| Eigenfunctions $y = \sin k\pi x$  |                                                          |     |

Now we go back to the matrix  $S$  and guess its eigenvectors. They come from  $\sin k\pi x$  at  $n$  points  $x = h, 2h, \dots, nh$ , equally spaced between 0 and 1. The spacing  $\Delta x$  is  $h = 1/(n+1)$ , so the  $(n+1)$ st point has  $(n+1)h = 1$ . Multiply that sine vector  $x$  by  $S$ :

|                                   |                                            |     |
|-----------------------------------|--------------------------------------------|-----|
| Eigenvalue of $S$ is positive     | $Sx = \lambda_k x = (2 - 2 \cos k\pi h) x$ | (7) |
| Eigenvector of $S$ is sine vector | $x = (\sin k\pi h, \dots, \sin nk\pi h)$ . |     |

## Problem Set 6.5

**Problems 1–13 are about tests for positive definiteness.**

- 1 Suppose the 2 by 2 tests  $a > 0$  and  $ac - b^2 > 0$  are passed. Then  $c > b^2/a > 0$ .

(i)  $\lambda_1$  and  $\lambda_2$  have the *same sign* because their product  $\lambda_1 \lambda_2$  equals \_\_\_\_.

(i) That sign is positive because  $\lambda_1 + \lambda_2$  equals \_\_\_\_.

*Conclusion:* The tests  $a > 0, ac - b^2 > 0$  guarantee positive eigenvalues  $\lambda_1, \lambda_2$ .

- 2 Which of  $S_1, S_2, S_3, S_4$  has two positive eigenvalues? Use a test, don't compute the  $\lambda$ 's. Also find an  $x$  so that  $x^T S_1 x < 0$ , so  $S_1$  is not positive definite.

$$S_1 = \begin{bmatrix} 5 & 6 \\ 6 & 7 \end{bmatrix} \quad S_2 = \begin{bmatrix} -1 & -2 \\ -2 & -5 \end{bmatrix} \quad S_3 = \begin{bmatrix} 1 & 10 \\ 10 & 100 \end{bmatrix} \quad S_4 = \begin{bmatrix} 1 & 10 \\ 10 & 101 \end{bmatrix}.$$

- 3 For which numbers  $b$  and  $c$  are these matrices positive definite?

$$S = \begin{bmatrix} 1 & b \\ b & 9 \end{bmatrix} \quad S = \begin{bmatrix} 2 & 4 \\ 4 & c \end{bmatrix} \quad S = \begin{bmatrix} c & b \\ b & c \end{bmatrix}.$$

With the pivots in  $D$  and multiplier in  $L$ , factor each  $A$  into  $LDL^T$ .

- 4 What is the function  $f = ax^2 + 2bxy + cy^2$  for each of these matrices? Complete the square to write each  $f$  as a sum of one or two squares  $f = d_1(\quad)^2 + d_2(\quad)^2$ .

$$S_1 = \begin{bmatrix} 1 & 2 \\ 2 & 9 \end{bmatrix} \quad S_2 = \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix} \quad f = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} S \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- 5 Write  $f(x, y) = x^2 + 4xy + 3y^2$  as a *difference* of squares and find a point  $(x, y)$  where  $f$  is negative. No minimum at  $(0, 0)$  even though  $f$  has positive coefficients.

- 6 The function  $f(x, y) = 2xy$  certainly has a saddle point and not a minimum at  $(0, 0)$ . What symmetric matrix  $S$  produces this  $f$ ? What are its eigenvalues?

- 7 Test to see if  $A^T A$  is positive definite in each case:  $A$  needs independent columns.

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

- 8 The function  $f(x, y) = 3(x + 2y)^2 + 4y^2$  is positive except at  $(0, 0)$ . What is the matrix in  $f = \begin{bmatrix} x & y \end{bmatrix} S \begin{bmatrix} x & y \end{bmatrix}^T$ ? Check that the pivots of  $A$  are 3 and 4.

- 9 Find the 3 by 3 matrix  $S$  and its pivots, rank, eigenvalues, and determinant:

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} S \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 4(x_1 - x_2 + 2x_3)^2.$$

- 10 Which 3 by 3 symmetric matrices  $S$  and  $T$  produce these quadratics?

$$\mathbf{x}^T S \mathbf{x} = 2(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3). \quad \text{Why is } S \text{ positive definite?}$$

$$\mathbf{x}^T T \mathbf{x} = 2(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_1x_3 - x_2x_3). \quad \text{Why is } T \text{ semidefinite?}$$

- 11 Compute the three upper left determinants of  $S$  to establish positive definiteness. Verify that their ratios give the second and third pivots.

$$\text{Pivots} = \text{ratios of determinants} \quad S = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 5 & 3 \\ 0 & 3 & 8 \end{bmatrix}.$$

- 12 For what numbers  $c$  and  $d$  are  $S$  and  $T$  positive definite? Test their 3 determinants:

$$S = \begin{bmatrix} c & 1 & 1 \\ 1 & c & 1 \\ 1 & 1 & c \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} 1 & 2 & 3 \\ 2 & d & 4 \\ 3 & 4 & 5 \end{bmatrix}.$$

- 13 Find a matrix with  $a > 0$  and  $c > 0$  and  $a + c > 2b$  that has a negative eigenvalue.

**Problems 14–20 are about applications of the tests.**

- 14 If  $S$  is positive definite then  $S^{-1}$  is positive definite. Best proof: The eigenvalues of  $S^{-1}$  are positive because \_\_\_\_\_. Second proof (only for 2 by 2):

$$\text{The entries of } S^{-1} = \frac{1}{ac - b^2} \begin{bmatrix} c & -b \\ -b & a \end{bmatrix} \text{ pass the determinant tests } \underline{\hspace{2cm}}.$$

- 15 If  $S$  and  $T$  are positive definite, their sum  $S + T$  is positive definite. Pivots and eigenvalues are not convenient for  $S + T$ . Better to use  $\mathbf{x}^T(S + T)\mathbf{x} > 0$ . Also  $S = A^T A$  and  $T = B^T B$  give  $S + T = \begin{bmatrix} A & B \end{bmatrix}^T \begin{bmatrix} A \\ B \end{bmatrix}$  with independent columns.

- 16 A positive definite matrix cannot have a zero (or even worse, a negative number) on its main diagonal. Show that this matrix fails to have  $\mathbf{x}^T S \mathbf{x} > 0$ :

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 4 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ is not positive when } (x_1, x_2, x_3) = ( \quad, \quad, \quad ).$$

- 17 A diagonal entry  $s_{jj}$  of a symmetric matrix cannot be smaller than all the  $\lambda$ 's. If it were, then  $S - s_{jj}I$  would have \_\_\_\_\_ eigenvalues and would be positive definite. But  $S - s_{jj}I$  has a \_\_\_\_\_ on the main diagonal.

- 18 If  $Sx = \lambda x$  then  $x^T Sx = \underline{\hspace{2cm}}$ . Why is this number positive when  $\lambda > 0$ ?
- 19 Reverse Problem 18 to show that *if all  $\lambda > 0$  then  $x^T Sx > 0$* . We must do this for **every nonzero  $x$** , not just the eigenvectors. So write  $x$  as a combination of the eigenvectors and *explain why all “cross terms” are  $x_i^T x_j = 0$* . Then  $x^T Sx$  is  
 $(c_1 x_1 + \cdots + c_n x_n)^T (c_1 \lambda_1 x_1 + \cdots + c_n \lambda_n x_n) = c_1^2 \lambda_1 x_1^T x_1 + \cdots + c_n^2 \lambda_n x_n^T x_n > 0$ .
- 20 Give a quick reason why each of these statements is true:
- (a) Every positive definite matrix is invertible.
  - (b) The only positive definite projection matrix is  $P = I$ .
  - (c) A diagonal matrix with positive diagonal entries is positive definite.
  - (d) A symmetric matrix with a positive determinant might not be positive definite!

**Problems 21–24 use the eigenvalues; Problems 25–27 are based on pivots.**

- 21 For which  $s$  and  $t$  do  $S$  and  $T$  have all  $\lambda > 0$  (therefore positive definite)?

$$S = \begin{bmatrix} s & -4 & -4 \\ -4 & s & -4 \\ -4 & -4 & s \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} t & 3 & 0 \\ 3 & t & 4 \\ 0 & 4 & t \end{bmatrix}.$$

- 22 From  $S = Q\Lambda Q^T$  compute the positive definite symmetric square root  $Q\sqrt{\Lambda}Q^T$  of each matrix. Check that this square root gives  $A^T A = S$ :

$$S = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix}.$$

- 23 You may have seen the equation for an ellipse as  $x^2/a^2 + y^2/b^2 = 1$ . What are  $a$  and  $b$  when the equation is written  $\lambda_1 x^2 + \lambda_2 y^2 = 1$ ? The ellipse  $9x^2 + 4y^2 = 1$  has axes with half-lengths  $a = \underline{\hspace{2cm}}$  and  $b = \underline{\hspace{2cm}}$ .
- 24 Draw the tilted ellipse  $x^2 + xy + y^2 = 1$  and find the half-lengths of its axes from the eigenvalues of the corresponding matrix  $S$ .
- 25 With positive pivots in  $D$ , the factorization  $S = LDL^T$  becomes  $L\sqrt{D}\sqrt{D}^T$ . (Square roots of the pivots give  $D = \sqrt{D}\sqrt{D}$ .) Then  $C = \sqrt{D}L^T$  yields the **Cholesky factorization**  $A = C^T C$  which is “symmetrized  $LU$ ”:

$$\text{From } C = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \text{ find } S. \quad \text{From } S = \begin{bmatrix} 4 & 8 \\ 8 & 25 \end{bmatrix} \text{ find } C = \mathbf{chol}(S).$$

- 26 In the Cholesky factorization  $S = C^T C$ , with  $C = \sqrt{D}L^T$ , the square roots of the pivots are on the diagonal of  $C$ . Find  $C$  (upper triangular) for

$$S = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 8 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 7 \end{bmatrix}.$$

- 27 The symmetric factorization  $S = LDL^T$  means that  $x^T S x = x^T L D L^T x$ :

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 0 \\ b/a & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & (ac - b^2)/a \end{bmatrix} \begin{bmatrix} 1 & b/a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The left side is  $ax^2 + 2bxy + cy^2$ . The right side is  $a(x + \frac{b}{a}y)^2 + \text{---}y^2$ .  
The second pivot completes the square! Test with  $a = 2, b = 4, c = 10$ .

- 28 Without multiplying  $S = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$ , find
- (a) the determinant of  $S$       (b) the eigenvalues of  $S$   
(c) the eigenvectors of  $S$       (d) a reason why  $S$  is symmetric positive definite.
- 29 For  $F_1(x, y) = \frac{1}{4}x^4 + x^2y + y^2$  and  $F_2(x, y) = x^3 + xy - x$  find the second derivative matrices  $S_1$  and  $S_2$ :

$$\text{Test for minimum} \quad S = \begin{bmatrix} \partial^2 F / \partial x^2 & \partial^2 F / \partial x \partial y \\ \partial^2 F / \partial y \partial x & \partial^2 F / \partial y^2 \end{bmatrix} \text{ is positive definite}$$

$S_1$  is positive definite so  $F_1$  is concave up (= convex). Find the minimum point of  $F_1$ .  
Find the saddle point of  $F_2$  (look only where first derivatives are zero).

- 30 The graph of  $z = x^2 + y^2$  is a bowl opening upward. The graph of  $z = x^2 - y^2$  is a saddle. The graph of  $z = -x^2 - y^2$  is a bowl opening downward. What is a test on  $a, b, c$  for  $z = ax^2 + 2bxy + cy^2$  to have a saddle point at  $(x, y) = (0, 0)$ ?
- 31 Which values of  $c$  give a bowl and which  $c$  give a saddle point for the graph of  $z = 4x^2 + 12xy + cy^2$ ? Describe this graph at the borderline value of  $c$ .

### The Minimum of a Function $F(x, y, z)$

What tests would you expect for a minimum point? First come zero slopes:

**First derivatives are zero**  $\frac{\partial F}{\partial x} = \frac{\partial F}{\partial y} = \frac{\partial F}{\partial z} = 0$  at the minimum point.

Next comes the linear algebra version of the usual calculus test  $d^2 f / dx^2 > 0$ :

**Second derivative matrix  $S$  is positive definite**  $S = \begin{bmatrix} F_{xx} & F_{xy} & F_{xz} \\ F_{yx} & F_{yy} & F_{yz} \\ F_{zx} & F_{zy} & F_{zz} \end{bmatrix}$

Here  $F_{xy} = \frac{\partial}{\partial x} \left( \frac{\partial F}{\partial y} \right) = \frac{\partial}{\partial y} \left( \frac{\partial F}{\partial x} \right) = F_{yx}$  is a 'mixed' second derivative.

### Challenge Problems

- 32** A *group* of nonsingular matrices includes  $AB$  and  $A^{-1}$  if it includes  $A$  and  $B$ . “Products and inverses stay in the group.” Which of these are groups (as in 2.7.37)? Invent a “subgroup” of two of these groups (not  $I$  by itself = the smallest group).
- (a) Positive definite symmetric matrices  $S$ .
  - (b) Orthogonal matrices  $Q$ .
  - (c) All exponentials  $e^{tA}$  of a fixed matrix  $A$ .
  - (d) Matrices  $P$  with positive eigenvalues.
  - (e) Matrices  $D$  with determinant 1.
- 33** When  $S$  and  $T$  are symmetric positive definite,  $ST$  might not even be symmetric. But its eigenvalues are still positive. Start from  $ST\mathbf{x} = \lambda\mathbf{x}$  and take dot products with  $T\mathbf{x}$ . Then prove  $\lambda > 0$ .
- 34** Write down the 5 by 5 sine matrix  $Q$  from Worked Example 6.5 C, containing the eigenvectors of  $S$  when  $n = 5$  and  $h = 1/6$ . Multiply  $SQ$  to see the five  $\lambda$ 's. The sum of  $\lambda$ 's should equal the trace 10. Their product should be  $\det S = 6$ .
- 35** Suppose  $C$  is positive definite (so  $\mathbf{y}^T C \mathbf{y} > 0$  whenever  $\mathbf{y} \neq \mathbf{0}$ ) and  $A$  has independent columns (so  $A\mathbf{x} \neq \mathbf{0}$  whenever  $\mathbf{x} \neq \mathbf{0}$ ). Apply the energy test to  $\mathbf{x}^T A^T C A \mathbf{x}$  to show that  $S = A^T C A$  is *positive definite: the crucial matrix in engineering*.
- 36** **Important!** Suppose  $S$  is positive definite with eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ .
- (a) What are the eigenvalues of the matrix  $\lambda_1 I - S$ ? Is it positive semidefinite?
  - (b) How does it follow that  $\lambda_1 \mathbf{x}^T \mathbf{x} \geq \mathbf{x}^T S \mathbf{x}$  for every  $\mathbf{x}$ ?
  - (c) Draw this conclusion: **The maximum value of  $\mathbf{x}^T S \mathbf{x} / \mathbf{x}^T \mathbf{x}$  is \_\_\_\_\_.**
- 37** For which  $a$  and  $c$  is this matrix positive definite? For which  $a$  and  $c$  is it positive semidefinite (this includes definite)?

$$S = \begin{bmatrix} a & a & a \\ a & a+c & a-c \\ a & a-c & a+c \end{bmatrix}$$

All 5 tests are possible.

The energy  $\mathbf{x}^T S \mathbf{x}$  equals

$$a(x_1 + x_2 + x_3)^2 + c(x_2 - x_3)^2.$$

## Table of Eigenvalues and Eigenvectors

How are the properties of a matrix reflected in its eigenvalues and eigenvectors? This question is fundamental throughout Chapter 6. A table that organizes the key facts may be helpful. Here are the special properties of the eigenvalues  $\lambda_i$  and the eigenvectors  $x_i$ .

|                                                        |                                                   |                                                                             |
|--------------------------------------------------------|---------------------------------------------------|-----------------------------------------------------------------------------|
| <b>Symmetric:</b> $S^T = S = Q\Lambda Q^T$             | real eigenvalues                                  | orthogonal $x_i^T x_j = 0$                                                  |
| <b>Orthogonal:</b> $Q^T = Q^{-1}$                      | all $ \lambda  = 1$                               | orthogonal $\bar{x}_i^T x_j = 0$                                            |
| <b>Skew-symmetric:</b> $A^T = -A$                      | imaginary $\lambda$ 's                            | orthogonal $\bar{x}_i^T x_j = 0$                                            |
| <b>Complex Hermitian:</b> $\bar{S}^T = S$              | real $\lambda$ 's                                 | orthogonal $\bar{x}_i^T x_j = 0$                                            |
| <b>Positive Definite:</b> $x^T S x > 0$                | all $\lambda > 0$                                 | orthogonal since $S^T = S$                                                  |
| <b>Markov:</b> $m_{ij} > 0, \sum_{i=1}^n m_{ij} = 1$   | $\lambda_{\max} = 1$                              | steady state $x > 0$                                                        |
| <b>Similar:</b> $A = BCB^{-1}$                         | $\lambda(A) = \lambda(C)$                         | $B$ times eigenvector of $C$                                                |
| <b>Projection:</b> $P = P^2 = P^T$                     | $\lambda = 1; 0$                                  | column space; nullspace                                                     |
| <b>Plane Rotation : cosine-sine</b>                    | $e^{i\theta}$ and $e^{-i\theta}$                  | $x = (1, i)$ and $(1, -i)$                                                  |
| <b>Reflection:</b> $I - 2uu^T$                         | $\lambda = -1; 1, \dots, 1$                       | $u$ ; whole plane $u^\perp$                                                 |
| <b>Rank One:</b> $uv^T$                                | $\lambda = v^T u; 0, \dots, 0$                    | $u$ ; whole plane $v^\perp$                                                 |
| <b>Inverse:</b> $A^{-1}$                               | $1/\lambda(A)$                                    | keep eigenvectors of $A$                                                    |
| <b>Shift:</b> $A + cI$                                 | $\lambda(A) + c$                                  | keep eigenvectors of $A$                                                    |
| <b>Stable Powers:</b> $A^n \rightarrow 0$              | all $ \lambda  < 1$                               | any eigenvectors                                                            |
| <b>Stable Exponential:</b> $e^{At} \rightarrow 0$      | all $\text{Re } \lambda < 0$                      | any eigenvectors                                                            |
| <b>Cyclic Permutation:</b> $P_{i,i+1} = 1, P_{n1} = 1$ | $\lambda_k = e^{2\pi i k/n} = \text{roots of } 1$ | $x_k = (1, \lambda_k, \dots, \lambda_k^{n-1})$                              |
| <b>Circulant:</b> $c_0 I + c_1 P + \dots$              | $\lambda_k = c_0 + c_1 e^{2\pi i k/n} + \dots$    | $x_k = (1, \lambda_k, \dots, \lambda_k^{n-1})$                              |
| <b>Tridiagonal:</b> $-1, 2, -1$ on diagonals           | $\lambda_k = 2 - 2 \cos \frac{k\pi}{n+1}$         | $x_k = \left( \sin \frac{k\pi}{n+1}, \sin \frac{2k\pi}{n+1}, \dots \right)$ |
| <b>Diagonalizable:</b> $A = X\Lambda X^{-1}$           | diagonal of $\Lambda$                             | columns of $X$ are independent                                              |
| <b>Schur:</b> $A = QTQ^{-1}$                           | diagonal of triangular $T$                        | columns of $Q$ if $A^T A = A A^T$                                           |
| <b>Jordan:</b> $A = BJB^{-1}$                          | diagonal of $J$                                   | each block gives 1 eigenvector                                              |
| <b>SVD:</b> $A = U\Sigma V^T$                          | $r$ singular values in $\Sigma$                   | eigenvectors of $A^T A, A A^T$ in $V, U$                                    |