

## Eigenvalues and eigenvectors

This topic applies to **square matrices** and the corresponding linear transformations, or to linear transformations  $T : V \rightarrow V$  (rather than the more general case  $T : V \rightarrow W$ ). We'll work with  $2 \times 2$  matrix transformations for now.

### 1. Diagonal matrices are easy

As you have seen in the past, diagonal matrices are especially easy. They represent problems in which the variables work separately. For example, let

$D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$ . Here are some ways in which  $D$  is easy:

- $D$  is the coefficient matrix of linear equations such as  $\begin{cases} 5x &= 8 \\ 3y &= 7 \end{cases}$ , easily solved separately.
- For  $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ , we have  $\tau_D(\mathbf{x}) = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5x \\ 3y \end{bmatrix}$ , a nonuniform scaling in which  $x$  and  $y$  are multiplied by separate constants.
- $D^2 = \begin{bmatrix} 25 & 0 \\ 0 & 9 \end{bmatrix}$ , obtained by treating each diagonal entry separately.

Figure 1 shows what  $\tau_D$  does to a unit square and vectors along the  $x$ - and  $y$ -axes.

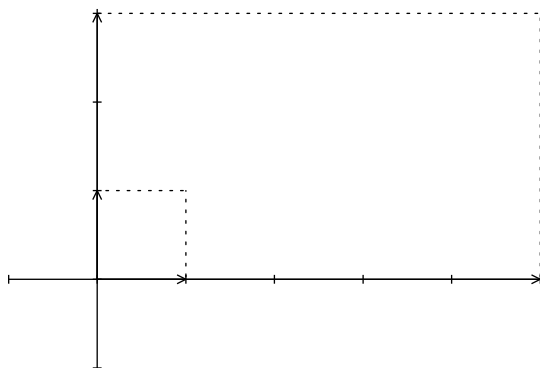


Figure 1: A diagonal transformation

**Problem U-1.** What, specifically, does  $\tau_D$  do to vectors along the  $x$ -axis? Along the  $y$ -axis? How about a vector pointing left along the  $x$ -axis?

## 2. Many linear transformations are secretly diagonal

Many linear transformations are like diagonal transformations if you look at them the right way—they are “diagonalizable”. If the matrix is not already diagonal, we need to choose a new basis, representing new axes.

Figure 2 shows the effects of the transformation  $\tau_A$  for  $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$ . But instead of showing the usual unit square and its image (a parallelogram), the figure shows a special slanted square and its image. Also shown are the vectors  $\mathbf{v}_1$ ,  $\mathbf{v}_2$  along the square, and their images.

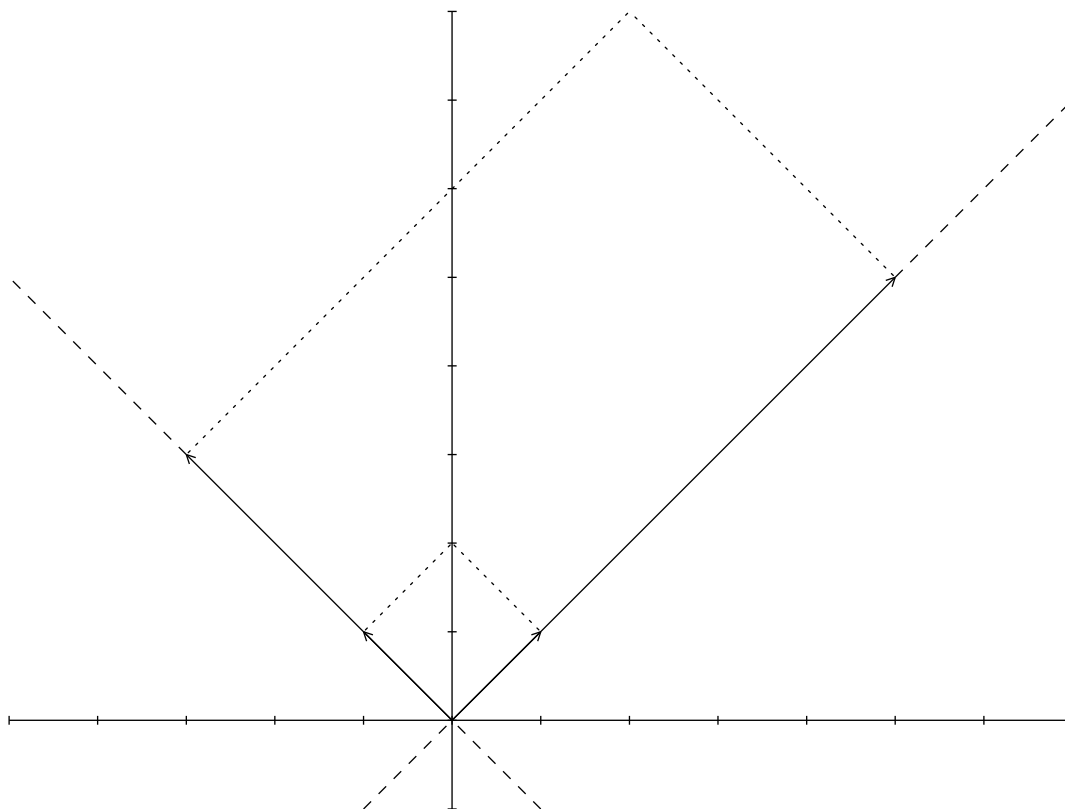


Figure 2: A diagonalizable transformation

**Problem U-2.** Find  $\tau_A(\mathbf{v}_1)$  and  $\tau_A(\mathbf{v}_2)$  numerically and see how they relate to  $\mathbf{v}_1$  and  $\mathbf{v}_2$ . Here  $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  and  $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ .

**Problem U-3.** Show that the matrix of  $\tau_A$  relative to the new basis  $\mathbf{v}_1$ ,  $\mathbf{v}_2$  is a diagonal matrix, specifically, the diagonal matrix  $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$ .

### 3. How to locate the new axes

The key to finding the new basis is to compare what happens to vectors on and off the desired axes. In Figures 3 and 4, what do you notice?

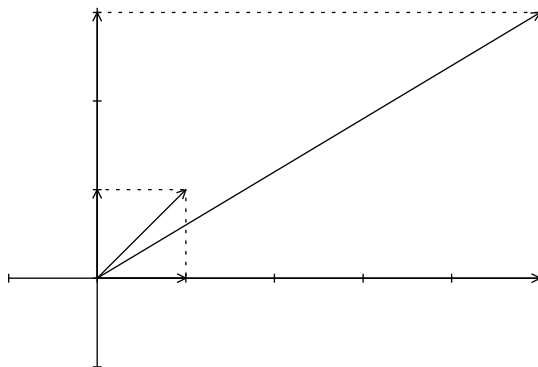


Figure 3: A diagonal transformation again

Notice that a vector *on* either of the desired axes has an image along the same line. In other words, it has been multiplied by a scalar. In contrast, for a vector *not* on one of the desired axes, its image is *not* along the same line.

This is the key! For any matrix transformation  $\tau_A$ , we look for vectors that get multiplied by a scalar. These tell us the new axes to use. As you will see, they are easy to compute.

**Definition.** A nonzero vector  $\mathbf{v}$  with  $\tau_A(\mathbf{v}) = \lambda\mathbf{v}$  for some scalar  $\lambda$  is called an *eigenvector* of  $\tau_A$ . The scalar  $\lambda$  is the corresponding *eigenvalue*<sup>1</sup>.

**Notes.**

1. It is traditional to use the Greek letter  $\lambda$  (lambda) for the scalar. This seems strange at first, but it's really helpful, in that when you see statements with  $\lambda$  you immediately realize they are about eigenvalues.
2. Figures 3 and 4 show  $T$  applied to just one vector not on the axes, but other vectors would behave similarly.
3. All we need is to choose *one* vector on each new axis for a new basis. Any nonzero vector along the axis will do.
4. The new axes are lines through the origin and so are subspaces. Therefore they are called *eigenspaces* for  $\tau_A$ . We'll emphasize eigenvectors

---

<sup>1</sup>The word “eigen” is German for “own” and has become traditional for these concepts. The “ei” is pronounced like “I”, as in “Einstein”. An alternative is to say “characteristic vector” and “characteristic value”.

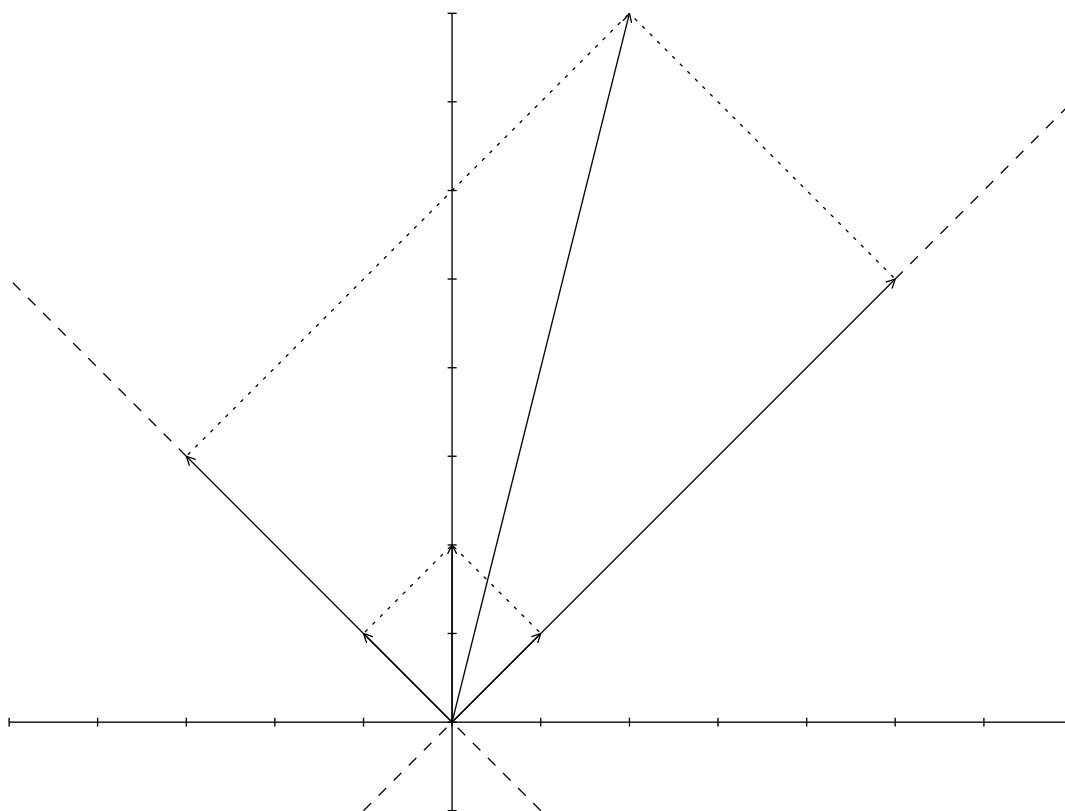


Figure 4: A diagonalizable transformation

for the present, but it's actually the eigenspaces that are the neater concept.

5. In general, for a  $2 \times 2$  matrix  $A$ , there can be either two, one, or no eigenspaces. If there are two, then you can make a new basis from two eigenvectors and get a diagonal matrix relative to the new basis. In this case,  $A$  is said to be *diagonalizable*.
6. In the more general setting  $T : V \rightarrow V$ , where  $V$  is a vector space, the concept is the same: If  $T(\mathbf{v}) = \lambda \mathbf{v}$  for a nonzero vector  $\mathbf{v}$ , then  $\mathbf{v}$  is an eigenvector and  $\lambda$  is an eigenvalue.
7. For a matrix transformation  $\tau_A$ , to say “ $\mathbf{v}$  is an eigenvector of  $A$ ” is the same thing as saying “ $\mathbf{v}$  is an eigenvector of  $\tau_A$ .” Obviously,  $\mathbf{v}$  is an eigenvector of  $A$  when  $\mathbf{v}$  is nonzero and  $A\mathbf{v} = \lambda \mathbf{v}$  for some scalar  $\lambda$ .
8. An eigenvalue could be negative, so that  $T(\mathbf{v})$  points in the opposite direction from  $\mathbf{v}$ . An eigenvalue can even be 0. (An *eigenvector* must be nonzero, though.)
9. There is still the question of how to *compute* eigenvectors and eigenvalues, even for  $2 \times 2$  matrices; this question will be answered below.

#### 4. Another example

The example in Figure 2 is especially neat, because the new axes are orthogonal (perpendicular to each other). This is not the situation in general. For the matrix  $A = \begin{bmatrix} 7 & -4 \\ 2 & 1 \end{bmatrix}$ , we do get two new axes (eigenspaces), but they are not orthogonal, as seen in Figure 5:

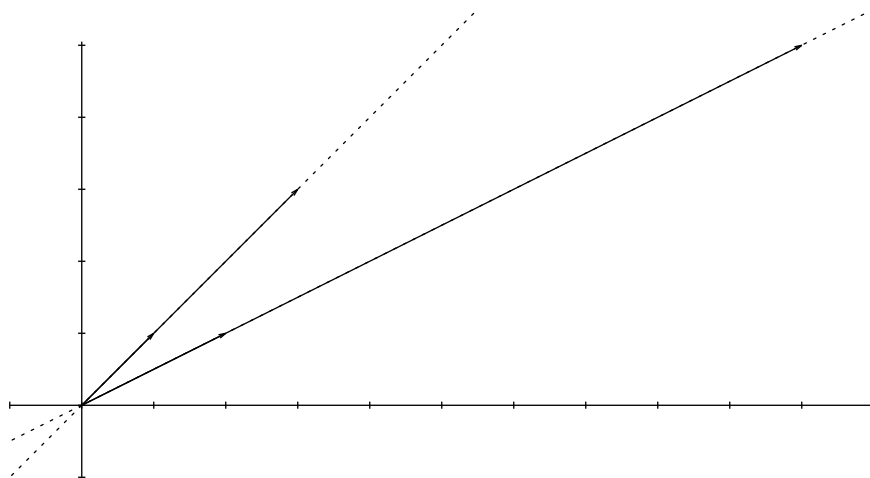


Figure 5: Eigenvectors in another example

**Problem U-4.** Check that  $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$  and  $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  are both eigenvectors of  $\tau_A$  for  $A = \begin{bmatrix} 7 & -4 \\ 2 & 1 \end{bmatrix}$ . What are the corresponding eigenvalues?

**Problem U-5.** Verify algebraically that  $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$  is not an eigenvector of  $A$ . (Check that  $\tau_A(\mathbf{v})$  is not a scalar times  $\mathbf{v}$ .)

**Problem U-6.** Explain why a rotation of  $\mathbb{R}^2$  by  $90^\circ$  can't have any eigenvectors. (Interestingly, if we use complex numbers, though, it does.)

## 5. How to find eigenvalues and eigenvectors

As a warmup, first recall that a square matrix  $A$  is “singular” if many equivalent conditions occur:  $\det A = 0$ ,  $A$  has no inverse,  $A$  is  $n \times n$  but has rank less than  $n$ , the nullspace of  $A$  has dimension  $> 0$  (so  $A\mathbf{v} = \mathbf{0}$  has a nonzero solution  $\mathbf{v}$ ), and so on. We'll need the first and last of these conditions. In other words, remember that  $A\mathbf{v} = \mathbf{0}$  has a nonzero solution  $\mathbf{v}$  if and only if  $\det A = 0$ .

Here is how to invent the method:

Given  $A$ , start with the equation  $A\mathbf{v} = \lambda\mathbf{v}$ . At this point, we don't yet know what the possible values of  $\lambda$  might be, or what nonzero vector  $\mathbf{v}$  might go with  $\lambda$ .

As in ordinary algebra, put everything on one side:  $A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0}$ .

Also as in ordinary algebra, it is tempting to try to factor out  $\mathbf{v}$ , but this doesn't work because  $(A - \lambda)$  doesn't make sense. You can't have a scalar minus a matrix!

A trick to get around this is to replace  $\lambda\mathbf{v}$  by  $\lambda I\mathbf{v}$ , where  $I$  is an identity matrix. If  $A$  is  $2 \times 2$ , for example, then  $\lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$ . So we can write  $A\mathbf{v} - \lambda I\mathbf{v} = \mathbf{0}$  and then get  $(A - \lambda I)\mathbf{v} = \mathbf{0}$ .

*Example:* For  $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$ ,  $A - \lambda I = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$   
 $= \begin{bmatrix} (4 - \lambda) & 1 \\ 1 & (4 - \lambda) \end{bmatrix}$ .

Next, from the warmup we notice that  $(A - \lambda I)\mathbf{v} = \mathbf{0}$  is possible for a nonzero vector  $\mathbf{v}$  if and only if  $\det(A - \lambda I) = 0$ . In the example, we see that  $\det(A - \lambda I) = (4 - \lambda)^2 - 1^2 = \lambda^2 - 8\lambda + 16 - 1 = \lambda^2 - 8\lambda + 15$ . We need to set this expression equal to 0 and solve.

In other words, the values of  $\lambda$  we are looking for are the *roots* of the *polynomial*  $\det(A - \lambda I)$ . This polynomial is called the *characteristic polynomial* of  $A$ , denoted here by  $p_A(\lambda)$ .

Find the roots using algebra:  $\lambda^2 - 8\lambda + 15 = 0$  factors to  $(\lambda - 5)(\lambda - 3) = 0$ , so  $\lambda = 5$  or  $\lambda = 3$ . Therefore the eigenvalues of  $A$  are 5 and 3 (in this example).

To find eigenvectors, consider each eigenvalue separately. For the eigenvalue 5, for example, we need to find a nonzero vector  $\mathbf{v}$  with  $A\mathbf{v} = 5\mathbf{v}$ , or equivalently,  $(A - 5I)\mathbf{v} = \mathbf{0}$ . This is just a set of linear equations!

In the example, for  $\lambda = 5$  we have  $A - 5I = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$ , so if we write  $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ ,  $(A - 5I)\mathbf{v} = \mathbf{0}$  becomes  $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ , or equivalently,  $\begin{cases} -x + y = 0 \\ x - y = 0 \end{cases}$ .

This set of equations is singular, with both equations giving the same information. Singular systems of equations may have seemed bad in the past, but here they are good, because we *do* want a nonzero solution. Also, it's no surprise, because we purposely chose  $\lambda$  so  $A - \lambda I$  is singular.

Choose any nonzero solution, say  $x = 1, y = 1$ . In other words, we have found that  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  is an eigenvector.

Similarly, for the eigenvalue  $\lambda = 3$ , we get  $A - 3I = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ , so  $(A - 3I)\mathbf{v} = \mathbf{0}$  becomes  $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ , or equivalently,  $\begin{cases} x + y = 0 \\ x + y = 0 \end{cases}$ . Let's choose the nonzero solution  $x = -1, y = 1$ . We have found that  $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$  is an eigenvector. ( $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$  would also do.)

To summarize the method, for given  $A$ :

*Step 1.* Find the characteristic polynomial of  $A$  by expanding  $\det(A - \lambda I)$ . (In the next section there is a shortcut for this when  $A$  is  $2 \times 2$ .)

*Step 2.* Set the characteristic polynomial  $= 0$  and find solutions (the roots of the characteristic polynomial). These are the eigenvalues of  $A$ .

*Step 3.* For each eigenvalue  $\lambda$ , treat  $(A - \lambda I)\mathbf{v} = \mathbf{0}$  as a set of linear equations and find a nonzero solution, which is a corresponding eigenvector.

---

<sup>1</sup>Some authors write  $\det(\lambda I - A)$  instead of  $\det(A - \lambda I)$  for the characteristic polynomial. This makes no difference if  $A$  is  $n \times n$  with  $n$  even, but it does make a difference in sign if  $n$  is odd.

## 6. More on the characteristic polynomial

For the  $2 \times 2$  case, let's see what happens if we “precompute” the characteristic polynomial using letters, by writing  $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ :

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

Notice that the constant term is  $\det A$ . The coefficient of  $\lambda$  is  $-(a + d)$ . The *trace* of a matrix means the sum of the diagonal entries, so the coefficient of  $\lambda$  is minus the trace.

Now you can see this fact:

*Observation.* If  $A$  is  $2 \times 2$ , then  $p_A(\lambda) = \lambda^2 - \text{trace}(A)\lambda + \det A$ .

Now it becomes easy to write down the characteristic polynomials in the examples from the early sections of this handout:

For  $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$  we get  $p_A(\lambda) = \lambda^2 - \text{trace}(A)\lambda + \det A = \lambda^2 - 8\lambda + 15$ .

For  $A = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$  we get  $p_A(\lambda) = \lambda^2 - \text{trace}(A)\lambda + \det A = \lambda^2 - 8\lambda + 15$ .

For  $A = \begin{bmatrix} 7 & -4 \\ 2 & 1 \end{bmatrix}$  we get  $p_A(\lambda) = \lambda^2 - \text{trace}(A)\lambda + \det A = \lambda^2 - 8\lambda + 15$ .

This explains why all three examples had the same eigenvalues: All three had the same characteristic polynomial.

**Problem U-7.** Invent two more  $2 \times 2$  matrices with characteristic polynomial  $\lambda^2 - 8\lambda + 15$ .

## 7. Easy facts

Notice that the characteristic polynomial is easy to find if the matrix  $A$  is diagonal or even triangular:

**Observation 1.** The eigenvalues of a triangular matrix are just the diagonal entries.

Why? For the case of an upper-triangular  $2 \times 2$  matrix,  $A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ , notice that  $\det(A - \lambda I) = \det \begin{bmatrix} (a - \lambda) & b \\ 0 & (d - \lambda) \end{bmatrix} = (\lambda - a)(\lambda - d)$ , so the roots are  $a$  and  $d$ . Larger matrices and lower-triangular matrices work the same.

Another easy fact:



**Observation 2.**  $A$  and  $A^t$  have the same characteristic polynomial.

The reason is that a matrix and its transpose have the same determinant, so  $\det(A - \lambda I)^t = \det(A - \lambda I)$ , or equivalently,  $\det(A - \lambda I^t) = \det(A - \lambda I)$ , or equivalently,  $p_{A^t}(\lambda) = p_A(\lambda)$ .

**Problem U-8.** A *stochastic matrix* is a square matrix such as  $A = \begin{bmatrix} .2 & .7 \\ .8 & .3 \end{bmatrix}$  in which the entries are  $\geq 0$  and each column has sum 1. Stochastic matrices have many applications, for example to shifts in populations over time. Let's stick to the  $2 \times 2$  case for convenience, but larger cases work the same.

- (a) Show that a  $2 \times 2$  matrix has *row* sums all equal to 1 if and only if  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  is an eigenvector for the eigenvalue 1.
- (b) Use Observation 2 to show that a stochastic  $2 \times 2$  matrix has the eigenvalue 1. This is the same thing as having a nonzero vector  $\mathbf{v}$  with  $T(\mathbf{v}) = \mathbf{v}$ , or in other words, a nonzero “fixed vector”.
- (c) Find a nonzero fixed vector for  $A$ .

**Problem U-9.** Explain: If  $A$  has the eigenvalue 0, then the eigenvectors for the eigenvalue 0 are all in the kernel of  $A$ , and  $A$  is singular.

## 8. A survey of $2 \times 2$ cases

So far we have not discussed whether *all*  $2 \times 2$  matrices are like the three examples—and they aren't! On the other hand, not many different kinds of things can happen. Here are the typical cases:

1. **Diagonal matrices.** There are two subcases:

- (a) **Scalar matrices**  $rI = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$ . Here the only eigenvalue is  $\lambda = r$  and every nonzero vector is an eigenvector.
- (b) **Diagonal matrices with distinct diagonal entries**  $\begin{bmatrix} d & 0 \\ 0 & e \end{bmatrix}$  with  $d \neq e$ . Here the eigenvectors lie on the  $x$ - and  $y$ -axes and the eigenvalues are just the diagonal entries.

2. **Symmetric matrices**, other than scalar matrices. In this case, there are two real eigenvalues and the eigenspaces (eigenvector directions) are orthogonal. See for instance the example in Figure 2.
3. **Matrices with two distinct real eigenvalues**, as in the example in Figure 5. Here there are two eigenspaces (lines of eigenvectors).

4. **Matrices with no real eigenvalues**, such as a  $90^\circ$  rotation. In this case there are no real eigenvectors either. However, if we are willing to use complex numbers then these matrices are like the previous case; they do have complex eigenvalues with corresponding eigenvectors.
5. **Matrices that have only one eigenvalue**, other than scalar matrices. Examples are (i) shears, such as the matrix  $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ ; (ii) nonzero matrices that are “nilpotent”, which means some power is the zero matrix, for example  $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ ; and (iii) nilpotent matrices added to scalar matrices, as in  $\begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ . These are strange in many ways. For example, they have only one line of eigenvectors.

Non-scalar  $2 \times 2$  matrices with only one eigenvalue are called “defective”.

You can begin to see how any  $2 \times 2$  matrix must be like one of these examples:

1. Find the characteristic polynomial, which has degree 2.
2. Factor it to get the roots (eigenvalues). You may need to use the quadratic formula, and the roots might be complex.
3. If the roots are real and distinct (different), then we get the two lines of eigenvectors in  $\mathbb{R}^2$ .
4. If the roots are identical (for example, if the characteristic polynomial is  $(\lambda - 3)^2$ ), then the matrix is either scalar or else is case 5.
5. Finally, if the roots are complex rather than real (in which case they are distinct), the matrix fits case 4.

All this is assuming that the matrix has real entries, but these cases are pretty much the same even for matrices with complex entries.

It turns out that even  $n \times n$  matrices split into pieces that fit these cases.

**Problem U-10.** Classify the eight examples in the handout with pictures of a house according to this scheme. (The first example is the identity matrix.)

**Problem U-11.** Show that a  $2 \times 2$  matrix is diagonalizable if and only if it is not defective. [This problem really should go after the next section. Also, “diagonalizable” here means possibly using complex numbers.]

## 9. Diagonalizing a square matrix

For a square matrix  $A$ , to *diagonalize*  $A$  means to find a square matrix  $P$  and diagonal matrix  $D$  with  $P^{-1}AP = D$ .

In this case we say that  $P$  *diagonalizes*  $A$ . Notice that  $P$  has to be nonsingular or it won't have an inverse.

Some matrices are “diagonalizable” and others are not.

*Example.* For  $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$ , it turns out that  $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$  is a good matrix to diagonalize  $A$ . In fact,  $P^{-1}AP = D$  for  $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$ .

Notice that  $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$  and  $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$  are eigenvectors of  $A$ . In fact,  $P$  is always made from eigenvectors:

*Proposition.* For  $n \times n$  matrices  $P$  and  $A$ ,  $P$  diagonalizes  $A \Leftrightarrow$  the columns of  $P$  are linearly independent eigenvectors of  $A$  and the diagonal entries of  $D$  are the corresponding eigenvalues.

(The linear independence of columns is needed in order for  $P$  to be nonsingular. It wouldn't work, for instance, to have all columns be the same. For the reasoning behind the Proposition, see below.)

The Proposition tells how to diagonalize  $A$ : Just find the eigenvalues and corresponding eigenvectors. If there are enough linearly independent eigenvectors, use them for the columns of  $P$ . The Example could be found that way.

Handy facts:

(1) Eigenvectors that belong to distinct eigenvalues are linearly independent. Therefore, if  $A$  is  $n \times n$  and has  $n$  distinct eigenvalues, then  $A$  is diagonalizable. (“Distinct” means “all different from one another”.)

(2) A real *symmetric* matrix can always be diagonalized, even if the eigenvalues are not all distinct (for example, if the characteristic polynomial factors as  $(\lambda - 1)(\lambda - 1)(\lambda - 4)(\lambda - 6)$ ), we would say the eigenvalues are 1, 1, 4, 6, not all distinct.

These two facts will be proved in class. Let's go back and prove the Proposition.

*Proof of the Proposition.* For “ $\Rightarrow$ ”: To say that  $P$  diagonalizes  $A$  means that  $P^{-1}AP = D$ . If we take this equation and multiply both sides by  $P$  on the left, we get  $AP = PD$ . Now let's try to interpret this second equation using eigenvalues and eigenvectors. We need two facts:

- First, whenever we multiply two matrices, such as  $AP$ , the columns of the matrix on the right work independently; if  $\mathbf{v}_1$  is the first column of  $P$  then  $A\mathbf{v}_1$  is the first column of  $AP$ , and so on.

For example,  $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & \cdot \\ 5 & \cdot \end{bmatrix} = \begin{bmatrix} 11 & \cdot \\ 24 & \cdot \end{bmatrix}$  and it doesn't matter what the other entries are.

- Second, whenever we multiply a matrix on the *right* by a diagonal matrix, as in  $PD$ , the effect is to multiply the columns of the left-hand matrix by scalars (the diagonal entries of  $D$ ), so that if  $\mathbf{v}_1$  is the first column of  $P$  then  $d_{11}\mathbf{v}_1$  is the first column of  $PD$ , and so on.

For example,  $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 10 & 0 \\ 0 & 100 \end{bmatrix} = \begin{bmatrix} 10 & 200 \\ 30 & 400 \end{bmatrix}$ .

Putting these two facts together, we see that  $A\mathbf{v}_1 = d_{11}\mathbf{v}_1$ , so  $\mathbf{v}_1$  is an eigenvector of  $A$  and  $d_{11}$  is the corresponding eigenvalue. (An eigenvector can't be  $\mathbf{0}$ , but obviously  $\mathbf{v}_1 \neq \mathbf{0}$  since it's a column of the invertible matrix  $P$ .) Now look at the second column  $\mathbf{v}_2$  of  $P$ ; by the same facts we now get  $A\mathbf{v}_2 = d_{22}\mathbf{v}_2$ , so  $\mathbf{v}_2$  is an eigenvector of  $P$  with corresponding eigenvalue  $d_{22}$ . The other columns of  $P$  work similarly. The Proposition also mentions that the columns of  $P$  are linearly independent, which is the same as saying that the column rank equals the number of columns, but this is automatic since  $P$  is invertible.

We still need to prove " $\Leftarrow$ " in the Proposition. We start by assuming we are given  $A$  and  $P$  such that the columns of  $P$  are eigenvectors of  $A$  and are linearly independent. The same two facts as before show that  $AP = PD$ . Since the columns of  $P$  are linearly independent,  $P$  has rank  $n$  and is invertible. Multiply both sides of  $AP = PD$  on the left by  $P^{-1}$ ; we get  $P^{-1}AP = D$ , as required.

**Problem U-12.** Diagonalize  $A = \begin{bmatrix} 7 & -4 \\ 2 & 1 \end{bmatrix}$ . (Use Problem U-4.)

**Problem U-13.** Explain how to interpret the equation  $P^{-1}AP = D$  in terms of a change of basis. What is the transformation? The old basis? The matrix relative to the old basis? The new basis? The transition matrix? The matrix relative to the new basis?

## 10. Eigenspaces

**Definition.** For a square matrix  $A$  and scalar  $\lambda$ , the *eigenspace* of  $\lambda$  consists of all  $\mathbf{v}$  with  $\tau_A(\mathbf{v}) = \lambda\mathbf{v}$ . We can call the eigenspace  $E_\lambda$ . In other words,  $E_\lambda$  consists of all eigenvectors for  $\lambda$  **and also the zero vector**.

One advantage of talking about eigenspaces instead of eigenvectors is that it's OK to talk about  $E_\lambda$  even when  $\lambda$  is not an eigenvalue of  $A$ ; in that case,  $E_\lambda = \{\mathbf{0}\}$ .

Another advantage is that it's easier to think about examples such as  $A = I$ , where  $E_1$  is the whole space, or  $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ , where  $E_2$  is a plane.

**Problem U-14.** Explain:  $E_\lambda$  is the nullspace of  $A - \lambda I$ , or equivalently, the kernel of  $\tau_{A-\lambda I}$ . Therefore  $E_\lambda$  is always a subspace.

**Problem U-15.** In each case, make up an example of a  $3 \times 3$  matrix so that  $E_5$  is (a)  $\{\mathbf{0}\}$ , (b) a line, (c) a plane, (d) all of  $\mathbb{R}^3$ . (Suggestion: Use diagonal matrices.)