6.1 Introduction to eigen value (*Refer 4.2 for applications)---:)https://github.com/AvinashShrivastav/MATHE

Chapter 6

Eigenvalues and Eigenvectors

6.1 Introduction to Eigenvalues

Linear equations $Ax = b$ come from steady state problems. Eigenvalues have their greatest importance in *dynamic problems*. The solution of $du/dt = Au$ is changing with time—growing or decaying or oscillating. We can't find it by elimination. This chapter enters a new part of linear algebra, based on $Ax = \lambda x$. All matrices in this chapter are square.

A good model comes from the powers $A, A^2, A^3, \ldots$ of a matrix. Suppose you need the hundredth power A^{100}. The starting matrix A becomes unrecognizable after a few steps, and A^{100} is very close to $[\,.6\ \ .6;\ \ .4\ \ .4\,]$:

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

A^{100} was found by using the *eigenvalues* of A, not by multiplying 100 matrices. Those eigenvalues (here they are 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 λ times the original x.

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

The eigenvalue λ 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 λ 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 λ'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 $\boldsymbol{\lambda} = \mathbf{1}$ and $\boldsymbol{\lambda} = \frac{1}{2}$. 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 $(1, -1)$:

$$x_1 = \begin{bmatrix} .6 \\ .4 \end{bmatrix} \quad \text{and} \quad Ax_1 = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}\begin{bmatrix} .6 \\ .4 \end{bmatrix} = x_1 \quad (Ax = x \text{ means that } \lambda_1 = 1)$$

$$x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad Ax_2 = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} .5 \\ -.5 \end{bmatrix} \quad (\text{this is } \tfrac{1}{2}x_2 \text{ so } \lambda_2 = \tfrac{1}{2}).$$

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.

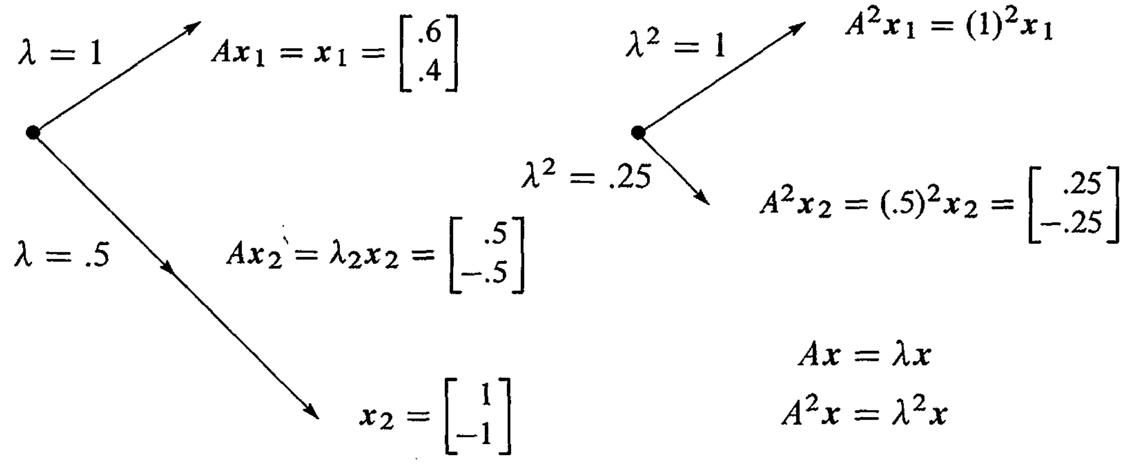


Figure 6.1: The eigenvectors keep their directions. A^2 has eigenvalues 1^2 and $(.5)^2$.

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

$$\textbf{Separate into eigenvectors} \qquad \begin{bmatrix} .8 \\ .2 \end{bmatrix} = x_1 + (.2)x_2 = \begin{bmatrix} .6 \\ .4 \end{bmatrix} + \begin{bmatrix} .2 \\ -.2 \end{bmatrix}. \tag{1}$$

Multiplying by A gives $(.7, .3)$, the first column of A^2. Do it separately for $\boldsymbol{x}_1$ and $(.2)\boldsymbol{x}_2$. Of course $A\boldsymbol{x}_1 = \boldsymbol{x}_1$. And A multiplies $\boldsymbol{x}_2$ by its eigenvalue $\frac{1}{2}$:

$$\textbf{Multiply each } \boldsymbol{x}_i \textbf{ by } \boldsymbol{\lambda}_i \qquad A\begin{bmatrix} .8 \\ .2 \end{bmatrix} = \begin{bmatrix} .7 \\ .3 \end{bmatrix} \quad \text{is} \quad \boldsymbol{x}_1 + \frac{1}{2}(.2)\boldsymbol{x}_2 = \begin{bmatrix} .6 \\ .4 \end{bmatrix} + \begin{bmatrix} .1 \\ -.1 \end{bmatrix}.$$

Each eigenvector is multiplied by its eigenvalue, when we multiply by A. We didn't need these eigenvectors to find A^2. But it is the good way to do 99 multiplications. At every step $\boldsymbol{x}_1$ is unchanged and $\boldsymbol{x}_2$ is multiplied by $(\frac{1}{2})$, so we have $(\frac{1}{2})^{99}$:

$$A^{99}\begin{bmatrix} .8 \\ .2 \end{bmatrix} \quad \text{is really} \quad \boldsymbol{x}_1 + (.2)(\tfrac{1}{2})^{99}\boldsymbol{x}_2 = \begin{bmatrix} .6 \\ .4 \end{bmatrix} + \begin{bmatrix} \text{very} \\ \text{small} \\ \text{vector} \end{bmatrix}.$$

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 $\boldsymbol{x}_1$ is a "steady state" that doesn't change (because $\lambda_1 = 1$). The eigenvector $\boldsymbol{x}_2$ is a "decaying mode" that virtually disappears (because $\lambda_2 = .5$). The higher the power of A, the closer its columns approach the steady state.

We mention that this particular A is a ***Markov matrix.*** Its entries are positive and every column adds to 1. Those facts guarantee that the largest eigenvalue is $\lambda = 1$ (as we found). Its eigenvector $\boldsymbol{x}_1 = (.6, .4)$ is the *steady state*—which all columns of A^k will approach. Section 8.3 shows how Markov matrices appear in applications like Google.

For projections we can spot the steady state ($\lambda = 1$) and the nullspace ($\lambda = 0$).

Example 2 **The projection matrix** $P = \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}$ **has eigenvalues** $\lambda = 1$ **and** $\lambda = 0$.

Its eigenvectors are $\boldsymbol{x}_1 = (1, 1)$ and $\boldsymbol{x}_2 = (1, -1)$. For those vectors, $P\boldsymbol{x}_1 = \boldsymbol{x}_1$ (steady state) and $P\boldsymbol{x}_2 = \boldsymbol{0}$ (nullspace). This example illustrates Markov matrices and singular matrices and (most important) symmetric matrices. All have special λ's and $\boldsymbol{x}$'s:

1. Each column of $P = \begin{bmatrix} .5 & .5 \\ .5 & .5 \end{bmatrix}$ 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 $P\boldsymbol{x} = 0\boldsymbol{x}$) fill up the nullspace. The eigenvectors for $\lambda = 1$ (which means $P\boldsymbol{x} = \boldsymbol{x}$) 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:

$$\textbf{Project each part} \quad \boldsymbol{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \end{bmatrix} \quad \textbf{projects onto} \quad P\boldsymbol{v} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \end{bmatrix}.$$

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).

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

Example 3 **The reflection matrix $R = \left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]$ 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*) $- I$:

$$\boldsymbol{R = 2P - I} \qquad \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} = 2\begin{bmatrix}.5 & .5\\.5 & .5\end{bmatrix} - \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}. \tag{2}$$

Here is the point. If $Px = \lambda x$ then $2Px = 2\lambda x$. The eigenvalues are doubled when the matrix is doubled. Now subtract $Ix = x$. The result is $(2P - I)x = (2\lambda - 1)x$. ***When a matrix is shifted by I, each λ is shifted by*** 1. No change in eigenvectors.

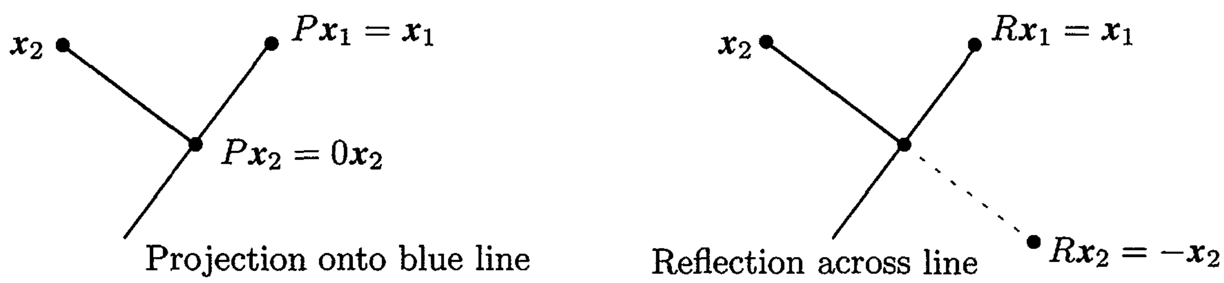


Figure 6.2: Projections P have eigenvalues 1 and 0. Reflections R have $\lambda = 1$ and -1. A typical x changes direction, but not the eigenvectors x_1 and x_2.

Key idea: The eigenvalues of R and P are related exactly as the matrices are related:

The eigenvalues of $R = 2P - I$ are $2(1) - 1 = 1$ and $2(0) - 1 = -1$.

The eigenvalues of R^2 are λ^2. In this case $R^2 = I$. Check $(1)^2 = 1$ and $(-1)^2 = 1$.

The Equation for the Eigenvalues

For projections and reflections we found λ's and x's by geometry: $Px = x$, $Px = \mathbf{0}$, $Rx = -x$. Now 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 λx to the left side. Write the equation $Ax = \lambda x$ as $(A - \lambda I)x = \mathbf{0}$. 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 λ, we find an eigenvector by solving $(A - \lambda I)x = \mathbf{0}$.

Eigenvalues first. If $(A - \lambda I)x = \mathbf{0}$ 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 λ:

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

$$\textbf{Equation for the eigenvalues} \qquad \det(A - \lambda I) = 0. \tag{3}$$

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

For each eigenvalue λ solve $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ or $A\boldsymbol{x} = \lambda\boldsymbol{x}$ to find an eigenvector $\boldsymbol{x}$.

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

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

$$\textit{\textbf{Subtract } } \boldsymbol{\lambda} \textit{\textbf{ from the diagonal to find}} \quad A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix}. \tag{4}$$

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 λ, is 2 times 2.

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

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

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

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

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

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

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: $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ is $A\boldsymbol{x} = \lambda\boldsymbol{x}$.

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, it is. The eigenvectors fill the nullspace: $A\boldsymbol{x} = 0\boldsymbol{x} = \mathbf{0}$. 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 λ subtracted along the diagonal, this determinant starts with λ^n or $-\lambda^n$. It is a polynomial in λ 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 λ, ***solve*** $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ ***to find an eigenvector*** $\boldsymbol{x}$.

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 $\lambda = 0$ and $\lambda = 5$:

$\lambda = 0$: 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 $\boldsymbol{x}$ is as good as $\boldsymbol{x}$. MATLAB's **eig**(A) divides by the length, to make the eigenvector into a unit vector.

We end with 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.) Similarly some n by n matrices don't have n independent eigenvectors. Without n eigenvectors, we don't have a basis. We can't write every $\boldsymbol{v}$ as a combination of eigenvectors. In the language of the next section, we can't diagonalize a matrix without n independent eigenvectors.

Good News, Bad News

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* λ'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* λ_1 *times* λ_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 on the main diagonal is called the ***trace*** of A:

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \textbf{\textit{trace}} = a_{11} + a_{22} + \cdots + a_{nn}. \tag{6}$$

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 λ's. But when the computation is wrong, they generally tell us so. To compute the correct λ's, go back to $\det(A - \lambda I) = 0$.

The determinant test makes the *product* of the λ's equal to the *product* of the pivots (assuming no row exchanges). But the sum of the λ's is not the sum of the pivots—as the example showed. The individual λ's have almost nothing to do with the pivots. In this new part of linear algebra, the key equation is really *nonlinear*: λ multiplies $\boldsymbol{x}$.

Why do the eigenvalues of a triangular matrix lie on 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* $\lambda = i$ *and* $\lambda = -i$. *Sum of* λ*'s* = trace = 0. *Product = determinant* = 1.

After a rotation, *no vector $Q\boldsymbol{x}$ stays in the same direction as $\boldsymbol{x}$* (except $\boldsymbol{x} = \mathbf{0}$ which is useless). There cannot be an eigenvector, unless we go to ***imaginary numbers***. Which we do.

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

Those λ'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:

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

Somehow these complex vectors $\boldsymbol{x}_1 = (1, i)$ and $\boldsymbol{x}_2 = (i, 1)$ 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 λ is $|\lambda| = 1$.

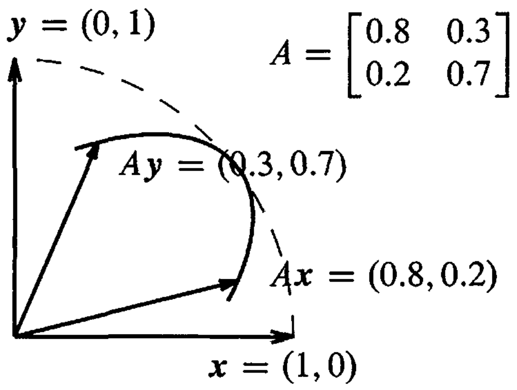
2. Q is a skew-symmetric matrix so each λ is pure imaginary.

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

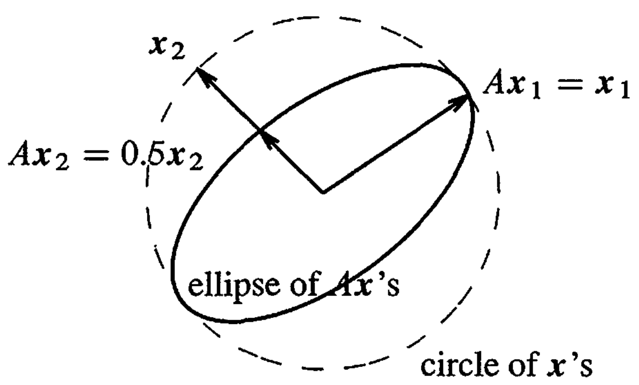
The eigenvectors for all these special matrices are perpendicular. Somehow $(i, 1)$ and $(1, i)$ are perpendicular (Chapter 10 explains the dot product of complex vectors).

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 $x = (1, 0)$. *The mouse makes this vector move around the unit circle.* At the same time the screen shows Ax, in color and also moving. Possibly Ax is ahead of x. Possibly Ax is behind x. *Sometimes Ax is parallel to x.* At that parallel moment, $Ax = \lambda x$ (at x_1 and x_2 in the second figure).



These are not eigenvectors



Ax lines up with x at eigenvectors

The eigenvalue λ is the length of Ax, when the unit eigenvector x lines up. The built-in choices for A illustrate three possibilities: 0, 1, or 2 directions where Ax crosses x.

0. There are *no real eigenvectors. Ax stays behind or ahead of x.* This means the eigenvalues and eigenvectors are complex, as they are for the rotation Q.

1. There is only *one* line of eigenvectors (unusual). The moving directions Ax and x touch but don't cross over. This happens for the last 2 by 2 matrix below.

2. There are eigenvectors in *two* independent directions. This is typical! Ax crosses x at the first eigenvector x_1, and it crosses back at the second eigenvector x_2. Then Ax and x cross again at $-x_1$ and $-x_2$.

You can mentally follow x and Ax for these five matrices. Under the matrices I will count their real eigenvectors. Can you see where Ax lines up with x?

$$A = \underset{\mathbf{2}}{\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}} \quad \underset{\mathbf{2}}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \quad \underset{\mathbf{0}}{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}} \quad \underset{\mathbf{1}}{\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}} \quad \underset{\mathbf{1}}{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}$$

When A is singular (rank one), its column space is a line. The vector Ax goes up and down that line while x circles around. One eigenvector x is along the line. Another eigenvector appears when $Ax_2 = \mathbf{0}$. Zero is an eigenvalue of a singular matrix.

■ 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 λ^2 and λ^{-1}, with the same eigenvectors.
4. The sum of the λ's equals the sum down the main diagonal of A (*the trace*). The product of the λ's equals the determinant.
5. Projections P, reflections R, 90° rotations Q have special eigenvalues $1, 0, -1, i, -i$. Singular matrices have $\lambda = 0$. Triangular matrices have λ's on their diagonal.

■ 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 $\lambda_1 + \lambda_2$ and the determinant $\lambda_1 \lambda_2$ for A and also A^2.

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

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

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 $\lambda_1 \lambda_2 = 3$. The eigenvectors come separately by solving $(A - \lambda I)x = \mathbf{0}$ which is $Ax = \lambda x$:

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

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

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

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

The trace of A^2 is $5 + 5$ which agrees with $1 + 9$. The determinant is $25 - 16 = 9$.

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 6.6). A is a *positive definite matrix* (Section 6.5) since $A = A^{\mathrm{T}}$ and the λ's are positive.

6.1 B Find the eigenvalues and eigenvectors of this 3 by 3 matrix A:

Symmetric matrix
Singular matrix
Trace $1 + 2 + 1 = 4$

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

Solution Since all rows of A add to zero, the vector $x = (1, 1, 1)$ gives $Ax = \mathbf{0}$. This is an eigenvector for the eigenvalue $\lambda = 0$. To find λ_2 and λ_3 I will compute the 3 by 3 determinant:

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

That factor $-\lambda$ confirms that $\lambda = 0$ is a root, and an eigenvalue of A. The other factors $(1 - \lambda)$ and $(3 - \lambda)$ give the other eigenvalues 1 and 3, adding to 4 (the trace). Each eigenvalue $0, 1, 3$ corresponds to an eigenvector:

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad Ax_1 = \mathbf{0}x_1 \qquad x_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad Ax_2 = \mathbf{1}x_2 \qquad x_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad Ax_3 = \mathbf{3}x_3\,.$$

I notice again that eigenvectors are perpendicular when A is symmetric.

The 3 by 3 matrix produced a third-degree (cubic) polynomial for $\det(A - \lambda I) = -\lambda^3 + 4\lambda^2 - 3\lambda$. We were lucky to find simple roots $\lambda = 0, 1, 3$. Normally we would use a command like **eig**(A), and the computation will never even use determinants (Section 9.3 shows a better way for large matrices).

The full command $[S, D] =$ **eig**(A) will produce unit eigenvectors in the columns of the **eigenvector matrix** S. The first one happens to have three minus signs, reversed from $(1, 1, 1)$ and divided by $\sqrt{3}$. The eigenvalues of A will be on the diagonal of the ***eigenvalue matrix*** (typed as D but soon called Λ).

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.

(a) Show from A how a row exchange can produce different eigenvalues.

(b) 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 λ_1 and λ_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 λ_1 and λ_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 $\boldsymbol{x}$ is an eigenvector, the way to find λ is to _____.

(b) If you know that λ is an eigenvalue, the way to find $\boldsymbol{x}$ is to _____.

9 What do you do to the equation $A\boldsymbol{x} = \lambda\boldsymbol{x}$, in order to prove (a), (b), and (c)?

(a) λ^2 is an eigenvalue of A^2, as in Problem 4.

(b) λ^{-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^∞. Explain from those answers why A^{100} is close to A^∞:

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

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 $\boldsymbol{x}_2$. Any idea why this should be?

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

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

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

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

(a) $P\boldsymbol{u} = \boldsymbol{u}$ comes from $(\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})\boldsymbol{u} = \boldsymbol{u}(\_\_\_\_\_)$. Then $\boldsymbol{u}$ is an eigenvector with $\lambda = 1$.

(b) If $\boldsymbol{v}$ is perpendicular to $\boldsymbol{u}$ show that $P\boldsymbol{v} = \boldsymbol{0}$. 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 $(Q - \lambda I)\boldsymbol{x} = \boldsymbol{0}$. Use $i^2 = -1$.

15 Every permutation matrix leaves $x = (1, 1, \ldots, 1)$ unchanged. Then $\lambda = 1$. Find two more λ'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 $\lambda = 0$:

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

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 $\lambda = (a + d + \sqrt{})/2$ and $\lambda = \underline{\qquad}$. Their sum is $\underline{\qquad}$. If A has $\lambda_1 = 3$ and $\lambda_2 = 4$ then $\det(A - \lambda I) = \underline{\qquad}$.

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 $B^{\mathrm{T}}B$

(c) the eigenvalues of $B^{\mathrm{T}}B$

(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 $\det(A^{\mathrm{T}} - \lambda I)$. That is true because $\underline{\qquad}$. 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 $M^{\mathrm{T}}(1, 1, 1) = (1, 1, 1)$. 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 λ'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 λ's and three eigenvectors:

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

25 Suppose A and B have the same eigenvalues $\lambda_1, \ldots, \lambda_n$ with the same independent eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. Then $A = B$. *Reason*: Any vector $\boldsymbol{x}$ is a combination $c_1\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{x}_n$. What is $A\boldsymbol{x}$? What is $B\boldsymbol{x}$?

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 λ'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:

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

6.2 Diagonalizing a Matrix

When x is an eigenvector, multiplication by A is just multiplication by a number λ: $Ax = \lambda x$. 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 Λ when we use the eigenvectors properly.*** This is the matrix form of our key idea. We start right off with that one essential computation.

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

Eigenvector matrix S
Eigenvalue matrix Λ
$$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}. \tag{1}$$

The matrix A is "diagonalized." We use capital lambda for the eigenvalue matrix, because of the small λ's (the eigenvalues) on its diagonal.

Proof Multiply A times its eigenvectors, which are the columns of S. The first column of AS is Ax_1. That is $\lambda_1 x_1$. Each column of S is multiplied by its eigenvalue λ_i:

A times S
$$AS = A\begin{bmatrix} & & \\ x_1 & \cdots & x_n \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ \lambda_1 x_1 & \cdots & \lambda_n x_n \\ & & \end{bmatrix}.$$

The trick is to split this matrix AS into S times Λ:

S times Λ
$$\begin{bmatrix} & & \\ \lambda_1 x_1 & \cdots & \lambda_n x_n \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ x_1 & \cdots & x_n \\ & & \end{bmatrix}\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = S\Lambda.$$

Keep those matrices in the right order! Then λ_1 multiplies the first column x_1, as shown. The diagonalization is complete, and we can write $AS = S\Lambda$ in two good ways:

$$AS = S\Lambda \quad \text{is} \quad S^{-1}AS = \Lambda \quad \text{or} \quad A = S\Lambda S^{-1}. \tag{2}$$

The matrix S 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 Λ have the same eigenvalues $\lambda_1, \ldots, \lambda_n$. The eigenvectors are different. The job of the original eigenvectors $x_1, \ldots, x_n$ was to diagonalize A. Those eigenvectors in S produce $A = S\Lambda S^{-1}$. You will soon see the simplicity and importance and meaning of the nth power $A^n = S\Lambda^n S^{-1}$.

Example 1 This A is triangular so the λ's are on the diagonal: $\lambda = 1$ and $\lambda = 6$.

$$\textbf{Eigenvectors} \quad \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix} \quad \underset{S^{-1}}{\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}} \underset{A}{\begin{bmatrix} 1 & 5 \\ 0 & 6 \end{bmatrix}} \underset{S}{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}} = \underset{\Lambda}{\begin{bmatrix} \mathbf{1} & 0 \\ 0 & \mathbf{6} \end{bmatrix}}$$

In other words $A = S\Lambda S^{-1}$. Then watch $A^2 = S\Lambda S^{-1}S\Lambda S^{-1}$. When you remove $S^{-1}S = I$, this becomes $S\Lambda^2 S^{-1}$. ***Same eigenvectors in S and squared eigenvalues in*** Λ^2.

The kth power will be $A^k = S\Lambda^k S^{-1}$ which is easy to compute:

$$\textbf{Powers of } A \qquad \begin{bmatrix} 1 & 5 \\ 0 & 6 \end{bmatrix}^k = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \\ & 6^k \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 6^k - 1 \\ 0 & 6^k \end{bmatrix}.$$

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 $A^2 = [1 \ \ 35; \ \ 0 \ \ 36]$ fits that formula when $k = 2$.

Here are four small remarks before we use Λ again.

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

Remark 2 ***We can multiply eigenvectors by any nonzero constants.*** $Ax = \lambda x$ will remain true. In Example 1, we can divide the eigenvector $(1, 1)$ by $\sqrt{2}$ to produce a unit vector.

Remark 3 The eigenvectors in S come in the same order as the eigenvalues in Λ. To reverse the order in Λ, put $(1, 1)$ before $(1, 0)$ in S:

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

To diagonalize A we *must* use an eigenvector matrix. From $S^{-1}AS = \Lambda$ we know that $AS = S\Lambda$. Suppose the first column of S is x. Then the first columns of AS and $S\Lambda$ 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:

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

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

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

There is no second eigenvector, so the 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 S).

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

Eigenvectors for n different λ's are independent. Then we can diagonalize A.

Independent $\boldsymbol{x}$ from different λ Eigenvectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j$ that correspond to distinct (all different) eigenvalues are linearly independent. An n by n matrix that has n different eigenvalues (no repeated λ's) must be diagonalizable.

Proof Suppose $c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 = \mathbf{0}$. Multiply by A to find $c_1\lambda_1\boldsymbol{x}_1 + c_2\lambda_2\boldsymbol{x}_2 = \mathbf{0}$. Multiply by λ_2 to find $c_1\lambda_2\boldsymbol{x}_1 + c_2\lambda_2\boldsymbol{x}_2 = \mathbf{0}$. Now subtract one from the other:

$$\text{Subtraction leaves} \quad (\lambda_1 - \lambda_2)c_1\boldsymbol{x}_1 = \mathbf{0}. \quad \text{Therefore } c_1 = 0.$$

Since the λ's are different and $\boldsymbol{x}_1 \neq \mathbf{0}$, we are forced to this conclusion that $c_1 = 0$. Similarly $c_2 = 0$. No other combination gives $c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 = \mathbf{0}$, so the eigenvectors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ must be independent.

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

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

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

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 $A = S\Lambda S^{-1}$ with those eigenvalues in the diagonal Λ:

$$\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix} = S\Lambda S^{-1}.$$

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

$$\textbf{Same } S \textbf{ for } A^2 \qquad A^2 = S\Lambda S^{-1}S\Lambda S^{-1} = S\Lambda^2 S^{-1}. \tag{4}$$

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

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

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

$$\textbf{Limit } k \to \infty \qquad A^\infty = \begin{bmatrix} .6 & 1 \\ .4 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ .4 & -.6 \end{bmatrix} = \begin{bmatrix} .6 & .6 \\ .4 & .4 \end{bmatrix}.$$

The limit has the eigenvector $\boldsymbol{x}_1$ in both columns. We saw this A^∞ on the very first page of the chapter. Now we see it coming, from powers like $A^{100} = S\Lambda^{100}S^{-1}$.

Question ***When does $A^k \to$ zero matrix?*** **Answer** ***All*** $|\lambda| < 1$.

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***:

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

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.

$$\text{Let } \boldsymbol{u}_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}. \text{ The rule } \begin{matrix} F_{k+2} = F_{k+1} + F_k \\ F_{k+1} = F_{k+1} \end{matrix} \text{ is } \boldsymbol{u}_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \boldsymbol{u}_k. \tag{5}$$

Every step multiplies by $A = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. After 100 steps we reach $\boldsymbol{u}_{100} = A^{100}\boldsymbol{u}_0$:

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

This problem is just right for eigenvalues. Subtract λ 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 $\left(-b \pm \sqrt{b^2 - 4ac}\right)/2a$:

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

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

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\lambda_1 - \lambda_2}\left(\begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} - \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}\right) \quad \text{or} \quad \boldsymbol{u}_0 = \frac{\boldsymbol{x}_1 - \boldsymbol{x}_2}{\lambda_1 - \lambda_2}. \tag{6}$$

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

100 steps from $\boldsymbol{u}_0$
$$\boldsymbol{u}_{100} = \frac{(\lambda_1)^{100}\boldsymbol{x}_1 - (\lambda_2)^{100}\boldsymbol{x}_2}{\lambda_1 - \lambda_2}. \tag{7}$$

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

$$F_{100} = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{100} - \left(\frac{1-\sqrt{5}}{2}\right)^{100}\right] \approx 3.54 \cdot 10^{20}. \tag{8}$$

Is this a whole number? *Yes.* The fractions and square roots must disappear, because Fibonacci's rule $F_{k+2} = F_{k+1} + F_k$ stays with integers. The second term in (8) is less than $\frac{1}{2}$, so it must move the first term to the nearest whole number:

$$k\text{th Fibonacci number} = \frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2} = \text{nearest integer to } \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^k. \tag{9}$$

The ratio of F_6 to F_5 is $8/5 = 1.6$. 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 $\boldsymbol{u}_{k+1} = A\mathbf{u}_k$. ***Each step multiplies by A.*** The solution is $\boldsymbol{u}_k = A^k\boldsymbol{u}_0$. We want to make clear how diagonalizing the matrix gives a quick way to compute A^k and find $\boldsymbol{u}_k$ in three steps.

The eigenvector matrix S produces $A = S\Lambda S^{-1}$. 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 S^{-1} multiplies S we get I***:

Powers of A
$$A^k\boldsymbol{u}_0 = (S\Lambda S^{-1})\cdots(S\Lambda S^{-1})\boldsymbol{u}_0 = S\Lambda^k S^{-1}\boldsymbol{u}_0$$

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

1. Write $\boldsymbol{u}_0$ as a combination $c_1\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{x}_n$ of the eigenvectors. Then $\boldsymbol{c} = S^{-1}\boldsymbol{u}_0$.

2. Multiply each eigenvector $\boldsymbol{x}_i$ by $(\lambda_i)^k$. Now we have $\Lambda^k S^{-1}\boldsymbol{u}_0$.

3. Add up the pieces $c_i(\lambda_i)^k\boldsymbol{x}_i$ to find the solution $\boldsymbol{u}_k = A^k\boldsymbol{u}_0$. This is $S\Lambda^k S^{-1}\boldsymbol{u}_0$.

Solution for $\boldsymbol{u}_{k+1} = A\boldsymbol{u}_k$
$$\boldsymbol{u}_k = A^k\boldsymbol{u}_0 = c_1(\lambda_1)^k\boldsymbol{x}_1 + \cdots + c_n(\lambda_n)^k\boldsymbol{x}_n. \tag{10}$$

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

$$\textbf{Step 1} \qquad \boldsymbol{u}_0 = \begin{bmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_n \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}. \quad \text{This says that } \boldsymbol{u}_0 = S\boldsymbol{c}. \tag{11}$$

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

$$A^k\boldsymbol{u}_0 = S\Lambda^kS^{-1}\boldsymbol{u}_0 = S\Lambda^k\boldsymbol{c} = \begin{bmatrix} \boldsymbol{x}_1 & \dots & \boldsymbol{x}_n \end{bmatrix} \begin{bmatrix} (\lambda_1)^k & & \\ & \ddots & \\ & & (\lambda_n)^k \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}. \tag{12}$$

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

Example 3 Start from $\boldsymbol{u}_0 = (1, 0)$. Compute $A^k\boldsymbol{u}_0$ when S and Λ contain these eigenvectors and eigenvalues:

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

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

Solution in three steps Find $\boldsymbol{u}_0 = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2$ and then $\boldsymbol{u}_k = c_1(\lambda_1)^k\boldsymbol{x}_1 + c_2(\lambda_2)^k\boldsymbol{x}_2$

Step 1 $\quad \boldsymbol{u}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{so} \quad c_1 = c_2 = \frac{1}{3}$

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

Step 3 $\quad$ Combine eigenvectors $c_1(\lambda_1)^k\boldsymbol{x}_1$ and $c_2(\lambda_2)^k\boldsymbol{x}_2$ into $\boldsymbol{u}_k$:

$$\boldsymbol{u}_k = A^k\boldsymbol{u}_0 \qquad \boldsymbol{u}_k = \frac{1}{3}2^k\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{1}{3}(-1)^k\begin{bmatrix} 1 \\ -1 \end{bmatrix}. \tag{13}$$

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 (powers λ^k will become $e^{\lambda t}$). Chapter 7 sees the same idea as "transforming to an eigenvector basis." The best example of all is a ***Fourier series***, built from the eigenvectors of d/dx.

Nondiagonalizable Matrices (Optional)

Suppose λ 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 λ 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 λ:

1. (Geometric Multiplicity = GM) Count the **independent eigenvectors** for λ. This is the dimension of the nullspace of $A - \lambda I$.

2. (Algebraic Multiplicity = AM) Count the **repetitions of λ** among the eigenvalues. Look at the n roots of $\det(A - \lambda I) = 0$.

If A has $\lambda = 4, 4, 4$, 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 $= 1$).

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

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

The vector called "repeats" in the Teaching Code **eigval** gives the algebraic multiplicity AM for each eigenvalue. When repeats = [1 1. . . 1] we know that the n eigenvalues are all different and A is diagonalizable. The sum of all components in "repeats" is always n, because every nth degree equation $\det(A - \lambda I) = 0$ has n roots (counting repetitions).

The diagonal matrix **D** in the Teaching Code **eigvec** gives the geometric multiplicity GM for each eigenvalue. This counts the independent eigenvectors. The total number of independent eigenvectors might be less than n. Then A is not diagonalizable.

We emphasize again: $\lambda = 0$ makes for easy computations, but these three matrices also have the same shortage of eigenvectors. Their repeated eigenvalue is $\lambda = 5$. Traces are 10, 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 AM $= 2$. But each $A - 5I$ has rank $r = 1$. The geometric multiplicity is GM $= 1$. There is only one line of eigenvectors for $\lambda = 5$, and these matrices are not diagonalizable.

Eigenvalues of AB and $A+B$

The first guess about the eigenvalues of AB is not true. An eigenvalue λ of A times an eigenvalue β of B usually does *not* give an eigenvalue of AB:

False proof $$ABx = A\beta x = \beta Ax = \beta\lambda x. \qquad (14)$$

It seems that β times λ 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:

> ***Commuting matrices share eigenvectors*** Suppose both A and B can be diagonalized. They share the same eigenvector matrix S 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). Then we cannot have $Px = \mathbf{0}$ at the same time as $Qx = \mathbf{0}$ (unless $x = \mathbf{0}$). If we knew the position exactly, we could not also know the momentum exactly. Problem 28 derives Heisenberg's uncertainty principle $\|Px\|\;\|Qx\| \geq \frac{1}{2}\|x\|^2$.

■ REVIEW OF THE KEY IDEAS ■

1. If A has n independent eigenvectors $x_1, \ldots, x_n$, they go into the columns of S.

   A is diagonalized by S $\qquad S^{-1}AS = \Lambda \quad \text{and} \quad A = S\Lambda S^{-1}.$

2. The powers of A are $A^k = S\Lambda^k S^{-1}$. The eigenvectors in S are unchanged.
3. The eigenvalues of A^k are $(\lambda_1)^k, \ldots, (\lambda_n)^k$ in the matrix Λ^k.
4. The solution to $u_{k+1} = Au_k$ starting from u_0 is $u_k = A^k u_0 = S\Lambda^k S^{-1} u_0$:

   $$u_k = c_1(\lambda_1)^k x_1 + \cdots + c_n(\lambda_n)^k x_n \quad \text{provided} \quad u_0 = c_1 x_1 + \cdots + c_n x_n.$$

   That shows Steps 1, 2, 3 (c's from $S^{-1}u_0$, λ^k from Λ^k, and x's from S)

5. A is diagonalizable if every eigenvalue has enough eigenvectors (GM = AM).

■ WORKED EXAMPLES ■

6.2 A The **Lucas numbers** are like the Fibonacci numbers except they start with $L_1 = 1$ and $L_2 = 3$. Following the 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 $\lambda_1^{100} + \lambda_2^{100}$.

Note The key point is that $\lambda_1 + \lambda_2 = 1$ and $\lambda_1^2 + \lambda_2^2 = 3$, when the λ's are $(1 \pm \sqrt{5})/2$. The Lucas number L_k is $\boldsymbol{\lambda_1^k + \lambda_2^k}$, since this is correct for L_1 and L_2.

Solution $\boldsymbol{u}_{k+1} = \left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]\boldsymbol{u}_k$ 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:

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

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

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

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

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

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

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

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 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 S that gives $S^{-1}AS = \Lambda$.

Solution What are the eigenvalues of the all-ones matrix **ones**(4)? 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 $\boldsymbol{x} = (1, 1, 1, 1)$ or (c, c, c, c). The other eigenvectors are perpendicular to $\boldsymbol{x}$ (since A is symmetric). The nicest eigenvector matrix S is the symmetric orthogonal Hadamard matrix H (normalized to unit column vectors):

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

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:

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

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

The determinant 125 counts the "spanning trees" in a graph with 5 nodes (all edges included). *Trees have no loops* (graphs and trees are in Section 8.2).

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

Problem Set 6.2

Questions 1–7 are about the eigenvalue and eigenvector matrices Λ and S.

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

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

(b) If $A = S\Lambda S^{-1}$ then $A^3 = (\quad)(\quad)(\quad)$ and $A^{-1} = (\quad)(\quad)(\quad)$.

2 If A has $\lambda_1 = 2$ with eigenvector $\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\lambda_2 = 5$ with $\boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, use $S\Lambda S^{-1}$ to find A. No other matrix has the same λ's and $\boldsymbol{x}$'s.

3 Suppose $A = S\Lambda S^{-1}$. What is the eigenvalue matrix for $A + 2I$? What is the eigenvector matrix? Check that $A + 2I = (\quad)(\quad)(\quad)^{-1}$.

4 True or false: If the columns of S (eigenvectors of A) are linearly independent, then

(a) A is invertible (b) A is diagonalizable

(c) S is invertible (d) S is diagonalizable.

5 If the eigenvectors of A are the columns of I, then A is a _____ matrix. If the eigenvector matrix S is triangular, then S^{-1} is triangular. Prove that A is also triangular.

6 Describe all matrices S 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 S^{-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 $S\Lambda^k S^{-1}\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ to find its second component. This is the kth 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:

$$\begin{aligned} G_{k+2} &= \tfrac{1}{2}G_{k+1} + \tfrac{1}{2}G_k \\ G_{k+1} &= G_{k+1} \end{aligned} \qquad \text{is} \qquad \begin{bmatrix} G_{k+2} \\ G_{k+1} \end{bmatrix} = \begin{bmatrix} & A & \end{bmatrix} \begin{bmatrix} G_{k+1} \\ G_k \end{bmatrix}.$$

(a) Find the eigenvalues and eigenvectors of A.

(b) Find the limit as $n \to \infty$ of the matrices $A^n = S\Lambda^n S^{-1}$.

(c) 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 $0, 1, 1, 2, 3, \ldots$ 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

(a) invertible (b) diagonalizable (c) not diagonalizable.

12 True or false: If the only eigenvectors of A are multiples of $(1, 4)$ then A has

(a) no inverse (b) a repeated eigenvalue (c) no diagonalization $S\Lambda S^{-1}$.

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 $A = \left[\begin{smallmatrix} 3 & 1 \\ 0 & 3 \end{smallmatrix}\right]$ 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 $A^k = S\Lambda^k S^{-1}$ approaches the zero matrix as $k \to \infty$ if and only if every λ has absolute value less than _____. Which of these matrices has $A^k \to 0$?

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

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

17 Find Λ and S 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 $S\Lambda^k S^{-1}$ 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 $S\Lambda^k S^{-1}$ 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 $A = S\Lambda S^{-1}$. 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 trace ST = trace TS, by adding the diagonal entries of ST and TS:

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

Choose T as ΛS^{-1}. Then $S\Lambda S^{-1}$ has the same trace as $\Lambda S^{-1} S = \Lambda$. The trace of A equals the trace of Λ = 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 $B = E^{\mathrm{T}}$ give

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

23 If $A = S\Lambda S^{-1}$, 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 S. Show that the A's form a subspace (cA and $A_1 + A_2$ have this same S). What is this subspace when $S = I$? 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 $A\boldsymbol{x} = \lambda\boldsymbol{x}$. If $\lambda = 0$ then $\boldsymbol{x}$ is in the nullspace. If $\lambda \neq 0$ then $\boldsymbol{x}$ 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 $R = S\sqrt{\Lambda}\, S^{-1}$. Why is there no real matrix square root of B?

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

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

Explain that last step by using the Schwarz inequality. Then Heisenberg's inequality says that $\|A\boldsymbol{x}\|/\|\boldsymbol{x}\|$ times $\|B\boldsymbol{x}\|/\|\boldsymbol{x}\|$ is at least $\frac{1}{2}$. It is impossible to get the position error and momentum error both very small.

29 If A and B have the same λ's with the same independent eigenvectors, their factorizations into _____ are the same. So $A = B$.

30 Suppose the same S diagonalizes both A and B. They have the same eigenvectors in $A = S\Lambda_1 S^{-1}$ and $B = S\Lambda_2 S^{-1}$. Prove that $AB = BA$.

31 (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) = $ *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 $\lambda^2 - \lambda - 1$.

6.4 Symmetric Matrices

For projection onto a plane in $\mathbf{R}^3$, the plane is full of eigenvectors (where $Px = x$). The other eigenvectors are *perpendicular* to the plane (where $Px = \mathbf{0}$). The eigenvalues $\lambda = 1, 1, 0$ are real. Three eigenvectors can be chosen perpendicular to each other. I have to write "can be chosen" because the two in the plane are not automatically perpendicular. This section makes that best possible choice for ***symmetric matrices***: ***The eigenvectors of*** $P = P^{\mathrm{T}}$ ***are perpendicular unit vectors.***

Now we open up to all symmetric matrices. It is no exaggeration to say that these 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 answer.

What is special about $Ax = \lambda x$ ***when*** A ***is symmetric***? We are looking for special properties of the eigenvalues λ and the eigenvectors x when $A = A^{\mathrm{T}}$.

The diagonalization $A = S\Lambda S^{-1}$ will reflect the symmetry of A. We get some hint by transposing to $A^{\mathrm{T}} = (S^{-1})^{\mathrm{T}}\Lambda S^{\mathrm{T}}$. Those are the same since $A = A^{\mathrm{T}}$. Possibly S^{-1} in the first form equals S^{T} in the second form. Then $S^{\mathrm{T}}S = I$. That makes each eigenvector in S orthogonal to the other eigenvectors. The key facts get first place in the Table at the end of this chapter, and here they are:

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

Those n orthonormal eigenvectors go into the columns of S. Every symmetric matrix can be diagonalized. ***Its eigenvector matrix*** S ***becomes an orthogonal matrix*** Q. Orthogonal matrices have $Q^{-1} = Q^{\mathrm{T}}$—what we suspected about S is true. To remember it we write $S = Q$, 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 $S\Lambda S^{-1}$ is in its special and particular form $Q\Lambda Q^{\mathrm{T}}$ for symmetric matrices:

(Spectral Theorem) Every symmetric matrix has the factorization $A = Q\Lambda Q^{\mathrm{T}}$ with real eigenvalues in Λ and orthonormal eigenvectors in $S = Q$:

Symmetric diagonalization $\quad A = Q\Lambda Q^{-1} = Q\Lambda Q^{\mathrm{T}} \quad \text{with} \quad Q^{-1} = Q^{\mathrm{T}}.$

It is easy to see that $Q\Lambda Q^{\mathrm{T}}$ is symmetric. Take its transpose. You get $(Q^{\mathrm{T}})^{\mathrm{T}}\Lambda^{\mathrm{T}}Q^{\mathrm{T}}$, which is $Q\Lambda Q^{\mathrm{T}}$ again. The harder part is to prove that every symmetric matrix has real λ'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 λ's in Λ and orthonormal $\boldsymbol{x}$'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 λ's and $\boldsymbol{x}$'s when $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ and $A - \lambda I = \begin{bmatrix} 1-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix}$.

Solution The determinant of $A - \lambda I$ is $\lambda^2 - 5\lambda$. The eigenvalues are 0 and 5 (*both real*). We can see them directly: $\lambda = 0$ is an eigenvalue because A is singular, and $\lambda = 5$ matches the *trace* down the diagonal of A: $0 + 5$ 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 into the columns of S (which is Q). Then $Q^{-1}AQ$ is Λ and $Q^{-1} = Q^{\mathrm{T}}$:

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

Now comes the n by n case. The λ's are real when $A = A^{\mathrm{T}}$ and $A\boldsymbol{x} = \lambda\boldsymbol{x}$.

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

Proof Suppose that $A\boldsymbol{x} = \lambda\boldsymbol{x}$. Until we know otherwise, λ might be a complex number $a + ib$ (a and b real). *Its complex conjugate is* $\overline{\lambda} = a - ib$. Similarly the components of $\boldsymbol{x}$ may be complex numbers, and switching the signs of their imaginary parts gives $\overline{\boldsymbol{x}}$. The good thing is that $\overline{\lambda}$ times $\overline{\boldsymbol{x}}$ is always the conjugate of λ times $\boldsymbol{x}$. So we can take conjugates of $A\boldsymbol{x} = \lambda\boldsymbol{x}$, remembering that A is real:

$$A\boldsymbol{x} = \lambda\boldsymbol{x} \quad \text{leads to} \quad A\overline{\boldsymbol{x}} = \overline{\lambda}\,\overline{\boldsymbol{x}}. \qquad \text{Transpose to} \quad \overline{\boldsymbol{x}}^{\mathrm{T}}A = \overline{\boldsymbol{x}}^{\mathrm{T}}\,\overline{\lambda}. \tag{1}$$

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

$$\overline{\boldsymbol{x}}^{\mathrm{T}}A\boldsymbol{x} = \overline{\boldsymbol{x}}^{\mathrm{T}}\lambda\boldsymbol{x} \qquad \text{and also} \qquad \overline{\boldsymbol{x}}^{\mathrm{T}}A\boldsymbol{x} = \overline{\boldsymbol{x}}^{\mathrm{T}}\,\overline{\lambda}\,\boldsymbol{x}. \tag{2}$$

The left sides are the same so the right sides are equal. One equation has λ, the other has $\overline{\lambda}$. They multiply $\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x} = |x_1|^2 + |x_2|^2 + \cdots =$ length squared which is not zero. *Therefore λ must equal $\overline{\lambda}$*, and $a + ib$ equals $a - ib$. The imaginary part is $b = 0$. Q.E.D.

The eigenvectors come from solving the real equation $(A - \lambda I)x = \mathbf{0}$. 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 λ's) are always perpendicular.

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

$$\textbf{Use } A^{\mathrm{T}} = A \qquad (\lambda_1 x)^{\mathrm{T}} y = (Ax)^{\mathrm{T}} y = x^{\mathrm{T}} A^{\mathrm{T}} y = x^{\mathrm{T}} A y = x^{\mathrm{T}} \lambda_2 y. \tag{3}$$

The left side is $x^{\mathrm{T}} \lambda_1 y$, the right side is $x^{\mathrm{T}} \lambda_2 y$. Since $\lambda_1 \neq \lambda_2$, this proves that $x^{\mathrm{T}} y = 0$. The eigenvector x (for λ_1) is perpendicular to the eigenvector y (for λ_2).

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

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

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

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

This is zero because $\lambda_1 + \lambda_2$ equals the trace $a + c$. Thus $x_1^{\mathrm{T}} x_2 = 0$. Eagle eyes might notice the special case $a = c$, $b = 0$ when $x_1 = x_2 = \mathbf{0}$. This case has repeated eigenvalues, as in $A = I$. It still has perpendicular eigenvectors $(1, 0)$ and $(0, 1)$.

This example shows the main goal of this section—***to diagonalize symmetric matrices A by orthogonal eigenvector matrices $S = Q$***. Look again at the result:

$$\textbf{Symmetry} \qquad A = S \Lambda S^{-1} \quad \text{becomes} \quad A = Q \Lambda Q^{\mathrm{T}} \quad \text{with} \quad Q^{\mathrm{T}} Q = I.$$

This says that every 2 by 2 symmetric matrix looks like

$$A = Q \Lambda Q^{\mathrm{T}} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} x_1^{\mathrm{T}} \\ x_2^{\mathrm{T}} \end{bmatrix}. \tag{5}$$

The columns x_1 and x_2 multiply the rows $\lambda_1 x_1^{\mathrm{T}}$ and $\lambda_2 x_2^{\mathrm{T}}$ to produce A:

$$\textbf{Sum of rank-one matrices} \qquad A = \lambda_1 x_1 x_1^{\mathrm{T}} + \lambda_2 x_2 x_2^{\mathrm{T}}. \tag{6}$$

This is the great factorization $Q \Lambda Q^{\mathrm{T}}$, written in terms of λ's and x's. When the symmetric matrix is n by n, there are n columns in Q multiplying n rows in Q^{T}. The n products $x_i x_i^{\mathrm{T}}$ are ***projection matrices***. Including the λ's, the spectral theorem $A = Q \Lambda Q^{\mathrm{T}}$ for symmetric matrices says that A is a combination of projection matrices:

$$A = \lambda_1 P_1 + \cdots + \lambda_n P_n \qquad \lambda_i = \text{eigenvalue}, \quad P_i = \text{projection onto eigenspace}.$$

Complex Eigenvalues of Real Matrices

Equation (1) went from $A\,x = \lambda\,x$ to $A\,\overline{x} = \overline{\lambda}\,\overline{x}$. In the end, λ and x were real. Those two equations were the same. But a *non*symmetric matrix can easily produce λ and x that are complex. In this case, $A\,\overline{x} = \overline{\lambda}\,\overline{x}$ is different from $A\,x = \lambda\,x$. It gives us a new eigenvalue (which is $\overline{\lambda}$) and a new eigenvector (which is $\overline{x}$):

For real matrices, complex λ's and x's come in "conjugate pairs."

$$\textit{If}\quad A\,x = \lambda\,x \quad \textit{then} \quad A\,\overline{x} = \overline{\lambda}\,\overline{x}.$$

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 λ and $\overline{\lambda}$. The eigenvectors must be x and $\overline{x}$, because A is real:

$$\begin{aligned} &\textbf{This is } \lambda\,x \qquad Ax = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = (\cos\theta + i\sin\theta)\begin{bmatrix} 1 \\ -i \end{bmatrix} \\ &\textbf{This is } \overline{\lambda}\,\overline{x} \qquad A\overline{x} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 \\ i \end{bmatrix} = (\cos\theta - i\sin\theta)\begin{bmatrix} 1 \\ i \end{bmatrix}. \end{aligned} \tag{7}$$

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.***

We apologize that a touch of complex numbers slipped in. They are unavoidable even when the matrix is real. Chapter 10 goes beyond complex numbers λ and complex vectors 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:

product of pivots = determinant = product of eigenvalues.

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

The number of positive eigenvalues of $A = A^{\mathrm{T}}$ equals the number of positive pivots.
Special case: A 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 A has one positive eigenvalue and one positive pivot:

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

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:

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

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 $A = A^{\mathrm{T}}$.

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

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

Watch the eigenvalues when L and L^{T} move toward the identity matrix: $A \to D$.

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 λ's move to the d's.

We repeat the proof for any $A = LDL^{\mathrm{T}}$. Move L toward I, by moving the off-diagonal entries to zero. The pivots are not changing and not zero. The eigenvalues λ 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. Q.E.D.

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 $A = A^{\mathrm{T}}$.***

Here is one idea for a proof. Change A slightly by a diagonal matrix $\mathbf{diag}(c, 2c, \ldots, nc)$. 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 $c \to 0$ we obtain n orthonormal eigenvectors of the original A—even if some eigenvalues of that A 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 matrices*, symmetric or not. This new factorization immediately produces $A = Q\Lambda Q^{\mathrm{T}}$ with a full set of real orthonormal eigenvectors when A is any symmetric matrix.

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

This is Schur's Theorem. We are looking for $AQ = QT$. The first column $\boldsymbol{q}_1$ of Q must be a unit eigenvector of A. Then the first columns of AQ and QT are $A\boldsymbol{q}_1$ and $t_{11}\boldsymbol{q}_1$. But the other columns of Q need not be eigenvectors when T is only triangular (not diagonal). So use any $n-1$ columns that complete $\boldsymbol{q}_1$ to a matrix Q_1 with orthonormal columns. At this point only the first columns of Q and T are set, where $A\boldsymbol{q}_1 = t_{11}\boldsymbol{q}_1$:

$$\overline{Q}_1^{\mathrm{T}} A Q_1 = \begin{bmatrix} \overline{\boldsymbol{q}}_1^{\mathrm{T}} \\ \vdots \\ \overline{\boldsymbol{q}}_n^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} A\boldsymbol{q}_1 & \cdots & A\boldsymbol{q}_n \end{bmatrix} = \begin{bmatrix} t_{11} & \cdots \\ 0 & \boxed{A_2} \\ \vdots & \\ 0 & \end{bmatrix}. \tag{8}$$

Now I will argue by "induction". *Assume Schur's factorization $A_2 = Q_2 T_2 Q_2^{-1}$ is possible for that matrix A_2 of size $n-1$.* Put the orthogonal (or unitary) matrix Q_2 and the triangular T_2 into the final Q and T:

$$Q = Q_1 \begin{bmatrix} 1 & 0 \\ 0 & Q_2 \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} t_{11} & \cdots \\ 0 & T_2 \end{bmatrix} \quad \text{and} \quad AQ = QT \quad \text{as desired.}$$

Note I had to allow $\boldsymbol{q}_1$ and Q_1 to be complex, in case A has complex eigenvalues. But if t_{11} is a real eigenvalue, then $\boldsymbol{q}_1$ and Q_1 can stay real. The induction step keeps everything real when A has real eigenvalues. Induction starts with 1 by 1, no problem.

Proof that T is the diagonal Λ when A is symmetric. Then we have $A = Q\Lambda Q^{\mathrm{T}}$.

Every symmetric A has real eigenvalues. Schur's $A = QTQ^{\mathrm{T}}$ with $Q^{\mathrm{T}}Q = I$ means that $T = Q^{\mathrm{T}}AQ$. This is a symmetric matrix (its transpose is $Q^{\mathrm{T}}AQ$). Now the key point: *If T is triangular and also symmetric, it must be diagonal*: $T = \Lambda$.

This proves $A = Q\Lambda Q^{\mathrm{T}}$. The matrix $A = A^{\mathrm{T}}$ has n orthonormal eigenvectors.

■ REVIEW OF THE KEY IDEAS ■

1. A symmetric matrix has *real eigenvalues* and *perpendicular eigenvectors*.
2. Diagonalization becomes $A = Q\Lambda Q^{\mathrm{T}}$ with an orthogonal 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 $A = A^{\mathrm{T}}$.
5. Every square matrix can be "triangularized" by $A = QTQ^{-1}$.

■ WORKED EXAMPLES ■

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

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

Solution All those properties can be predicted! With real eigenvalues in Λ and orthonormal eigenvectors in Q, the matrix $A = Q\Lambda Q^{\mathrm{T}}$ 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 are positive and negative like the eigenvalues, since A is symmetric.

The matrix must be a reflection. Vectors in the direction of $\boldsymbol{x}_1$ are unchanged by A (since $\lambda = 1$). Vectors in the perpendicular direction are reversed (since $\lambda = -1$). The reflection $A = Q\Lambda Q^{\mathrm{T}}$ is across the "θ-line". Write c for $\cos\theta$, 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 $\boldsymbol{x} = (1, 0)$ goes to $A\boldsymbol{x} = (\cos 2\theta, \sin 2\theta)$ on the 2θ-line. And $(\cos 2\theta, \sin 2\theta)$ goes back across the θ-line to $\boldsymbol{x} = (1, 0)$.

6.4 B Find the eigenvalues of A_3 and B_4, and check the orthogonality of their first two eigenvectors. Graph these eigenvectors to see discrete sines and cosines:

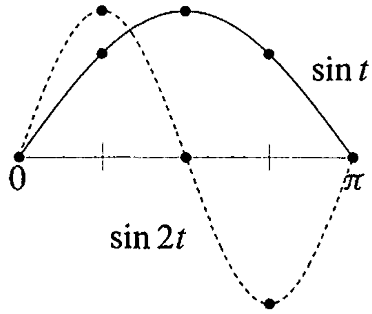
$$A_3 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \qquad B_4 = \begin{bmatrix} 1 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 1 \end{bmatrix}$$

The $-1, 2, -1$ pattern in both matrices is a "second difference". Section 8.1 will explain how this is like a second derivative. Then $A\boldsymbol{x} = \lambda\boldsymbol{x}$ and $B\boldsymbol{x} = \lambda\boldsymbol{x}$ 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. The matrices 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 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 S gives the "Discrete Sine Transform" and the graph shows how the first two eigenvectors fall onto sine curves. Please draw the third eigenvector onto a third sine curve!

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

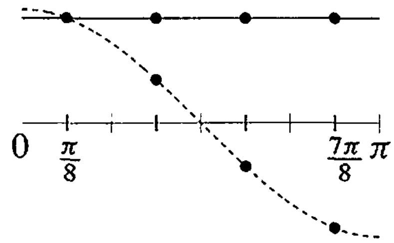
Eigenvector matrix for A_3



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" and the graph shows how the first two eigenvectors fall onto cosine curves. (Please plot the third eigenvector.) These eigenvectors match cosines at the *halfway points* $\frac{\pi}{8}, \frac{3\pi}{8}, \frac{5\pi}{8}, \frac{7\pi}{8}$.

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \sqrt{2}-1 & -1 & 1-\sqrt{2} \\ 1 & 1-\sqrt{2} & -1 & \sqrt{2}-1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

Eigenvector matrix for B_4



S and C have orthogonal columns (eigenvectors of the symmetric A_3 and B_4). When we multiply a vector by S or C, that signal splits into pure frequencies—as a musical chord separates into pure notes. This is the most useful and insightful transform in all of signal processing. Here is a MATLAB code to create B_8 and its eigenvector matrix C:

```
n=8; e=ones(n-1,1); B=2*eye(n)-diag(e,-1)-diag(e,1); B(1,1)=1; B(n,n)=1;
[C, Λ] = eig(B);
plot(C(:,1:4), '-o')
```

Problem Set 6.4

1 Write A as $M + N$, symmetric matrix plus skew-symmetric matrix:

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 4 & 3 & 0 \\ 8 & 6 & 5 \end{bmatrix} = M + N \qquad (M^{\mathrm{T}} = M, N^{\mathrm{T}} = -N).$$

For any square matrix, $M = \frac{A+A^{\mathrm{T}}}{2}$ and $N =$ ____ add up to A.

2 If C is symmetric prove that $A^{\mathrm{T}}CA$ is also symmetric. (Transpose it.) When A is 6 by 3, what are the shapes of C and $A^{\mathrm{T}}CA$?

3 Find the eigenvalues and the unit eigenvectors of

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

4 Find an orthogonal matrix Q that diagonalizes $A = \begin{bmatrix} -2 & 6 \\ 6 & 7 \end{bmatrix}$. What is Λ?

5 Find an orthogonal matrix Q that diagonalizes this symmetric matrix:

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

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

7 (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?

8 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 zero.

9 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 $A\bar{x} = \bar{\lambda}\bar{x}$.) Prove that every real 3 by 3 matrix has at least one real eigenvalue.

10 Here is a quick "proof" that the eigenvalues of all real matrices are real:

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

Find the flaw in this reasoning—a hidden assumption that is not justified. You could test those steps on the 90° rotation matrix $[0 \;\; -1; \;\; 1 \;\; 0]$ with $\lambda = i$ and $x = (i, 1)$.

11 Write A and B in the form $\lambda_1 x_1 x_1^{\mathrm{T}} + \lambda_2 x_2 x_2^{\mathrm{T}}$ of the spectral theorem $Q\Lambda Q^T$:

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

12 Every 2 by 2 symmetric matrix is $\lambda_1 x_1 x_1^{\mathrm{T}} + \lambda_2 x_2 x_2^{\mathrm{T}} = \lambda_1 P_1 + \lambda_2 P_2$. Explain $P_1 + P_2 = x_1 x_1^{\mathrm{T}} + x_2 x_2^{\mathrm{T}} = I$ from columns times rows of Q. Why is $P_1 P_2 = 0$?

13 What are the eigenvalues of $A = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$? Create a 4 by 4 skew-symmetric matrix ($A^{\mathrm{T}} = -A$) and verify that all its eigenvalues are imaginary.

14 (Recommended) This matrix M is skew-symmetric 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:

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

15 Show that A (**symmetric but complex**) has only one line of eigenvectors:

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

$A^{\mathrm{T}} = A$ is not such a special property for complex matrices. The good property is $\overline{A}^{\mathrm{T}} = A$ (Section 10.2). Then all λ's are real and eigenvectors are orthogonal.

16 Even if A is rectangular, the block matrix $B = \begin{bmatrix} 0 & A \\ A^{\mathrm{T}} & 0 \end{bmatrix}$ is symmetric:

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

(a) Show that $-\lambda$ is also an eigenvalue, with the eigenvector $(y, -z)$.

(b) Show that $A^{\mathrm{T}}Az = \lambda^2 z$, so that λ^2 is an eigenvalue of $A^{\mathrm{T}}A$.

(c) If $A = I$ (2 by 2) find all four eigenvalues and eigenvectors of B.

17 If $A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in Problem 16, find all three eigenvalues and eigenvectors of B.

18 ***Another proof that eigenvectors are perpendicular when $A = A^{\mathrm{T}}$.*** Two steps:

1. Suppose $Ax = \lambda x$ and $Ay = 0y$ 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 $Ay = \beta y$, apply this argument to $A - \beta I$. The eigenvalue of $A - \beta I$ moves to zero and the eigenvectors stay the same—so they are perpendicular.

19 Find the eigenvector matrix S for A and for B. Show that S doesn't collapse at $d = 1$, even though $\lambda = 1$ is repeated. Are the eigenvectors perpendicular?

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

20 Write a 2 by 2 *complex* matrix with $\overline{A}^{\mathrm{T}} = A$ (a "Hermitian matrix"). Find λ_1 and λ_2 for your complex matrix. Adjust equations (1) and (2) to show that *the eigenvalues of a Hermitian matrix are real.*

21 ***True*** (with reason) or ***false*** (with example). "Orthonormal" is not assumed.

(a) A matrix with real eigenvalues and eigenvectors is symmetric.

(b) A matrix with real eigenvalues and orthogonal eigenvectors is symmetric.

(c) The inverse of a symmetric matrix is symmetric.

(d) The eigenvector matrix S of a symmetric matrix is symmetric.

22 (A paradox for instructors) If $AA^{\mathrm{T}} = A^{\mathrm{T}}A$ 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: They must have the same S and Λ. Therefore A equals A^{T}.

23 (Recommended) Which of these classes of matrices do A and B belong to: Invertible, orthogonal, projection, permutation, diagonalizable, Markov?

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

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

24 What number b in $\left[\begin{smallmatrix} 2 & b \\ 1 & 0 \end{smallmatrix}\right]$ makes $A = Q\Lambda Q^{\mathrm{T}}$ possible? What number makes $A = S\Lambda S^{-1}$ impossible? What number makes A^{-1} impossible?

25 Find all 2 by 2 matrices that are orthogonal and also symmetric. Which two numbers can be eigenvalues?

26 This A is nearly symmetric. But its eigenvectors are far from orthogonal:

$$A = \begin{bmatrix} 1 & 10^{-15} \\ 0 & 1 + 10^{-15} \end{bmatrix} \quad \text{has eigenvectors} \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad [?]$$

What is the angle between the eigenvectors?

27 (MATLAB) Take two symmetric matrices with different eigenvectors, say $A = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right]$ and $B = \left[\begin{smallmatrix} 8 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. 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 λ_1 and λ_2 appear to be on a collision course at certain values of t. "Yet at the last minute they turn aside." How close do they come?

Challenge Problems

28 For complex matrices, the symmetry $A^{\mathrm{T}} = A$ that produces real eigenvalues changes to $\overline{A}^{\mathrm{T}} = A$. From $\det(A - \lambda I) = 0$, find the eigenvalues of the 2 by 2 "Hermitian" matrix $A = [4 \;\; 2+i; \;\; 2-i \;\; 0] = \overline{A}^{\mathrm{T}}$. To see why eigenvalues are real when $\overline{A}^{\mathrm{T}} = A$, adjust equation (1) of the text to $\overline{A}\,\overline{x} = \overline{\lambda}\,\overline{x}$.

Transpose to $\overline{x}^{\mathrm{T}}\overline{A}^{\mathrm{T}} = \overline{x}^{\mathrm{T}}\overline{\lambda}$. **With** $\overline{A}^{\mathrm{T}} = A$, **reach equation (2):** $\lambda = \overline{\lambda}$.

Example 13

Theorem 5.28 shows the operator on $\mathbb{R}^4$ in Example 12 is not diagonalizable because the geometric multiplicity of $\lambda_1 = 3$ is 2, while its algebraic multiplicity is 3.

Example 14

Let $L: \mathbb{R}^3 \to \mathbb{R}^3$ be a rotation about the z-axis through an angle of $\frac{\pi}{3}$. Then the matrix for L with respect to the standard basis is

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

as described in Table 5.1. Using $\mathbf{A}$, we calculate $p_L(x) = x^3 - 2x^2 + 2x - 1 = (x-1)(x^2 - x + 1)$, where the quadratic factor has no real roots. Therefore, $\lambda = 1$ is the only eigenvalue, and its algebraic multiplicity is 1. Hence, by Theorem 5.28, L is not diagonalizable because the sum of the algebraic multiplicities of its eigenvalues equals 1, which is less than $\dim(\mathbb{R}^3) = 3$.

The Cayley-Hamilton Theorem

We conclude this section with an interesting relationship between a matrix and its characteristic polynomial. If $p(x) = a_n x^n + a_{n-1}x^{n-1} \cdots + a_1 x + a_0$ is any polynomial and $\mathbf{A}$ is an $n \times n$ matrix, we define $p(\mathbf{A})$ to be the $n \times n$ matrix given by $p(\mathbf{A}) = a_n \mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} \cdots + a_1 \mathbf{A} + a_0 \mathbf{I}_n$.

Theorem 5.29 (Cayley-Hamilton Theorem) Let $\mathbf{A}$ be an $n \times n$ matrix, and let $p_{\mathbf{A}}(x)$ be its characteristic polynomial. Then $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}_n$.

The Cayley-Hamilton Theorem is an important result in advanced linear algebra. We have placed its proof in Appendix A for the interested reader.

Example 15

Let $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 4 & -1 \end{bmatrix}$. Then $p_{\mathbf{A}}(x) = x^2 - 2x - 11$ (verify!). The Cayley-Hamilton Theorem states that $p_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}_2$. To check this, note that

$$p_{\mathbf{A}}(\mathbf{A}) = \mathbf{A}^2 - 2\mathbf{A} - 11\mathbf{I}_2 = \begin{bmatrix} 17 & 4 \\ 8 & 9 \end{bmatrix} - \begin{bmatrix} 6 & 4 \\ 8 & -2 \end{bmatrix} - \begin{bmatrix} 11 & 0 \\ 0 & 11 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

6.5 Positive Definite Matrices

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 these matrices. You may say, just find the eigenvalues and test $\lambda > 0$. That is exactly what we want to avoid. Calculating eigenvalues is work. When the λ's are needed, we can compute them. But if we just want to know that they 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.

The λ's are automatically real because the matrix is symmetric.

Start with 2 by 2. When does $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ ***have*** $\lambda_1 > 0$ ***and*** $\lambda_2 > 0$?

The eigenvalues of A are positive if and only if $a > 0$ ***and*** $ac - b^2 > 0$.

$$A_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \text{ is } \textit{not} \text{ positive definite because } ac - b^2 = 1 - 4 < 0$$

$$A_2 = \begin{bmatrix} 1 & -2 \\ -2 & 6 \end{bmatrix} \text{ is positive definite because } a = 1 \text{ and } ac - b^2 = 6 - 4 > 0$$

$$A_3 = \begin{bmatrix} -1 & 2 \\ 2 & -6 \end{bmatrix} \text{ is } \textit{not} \text{ positive definite (even with } \det A = +2) \text{ because } a = -1$$

Notice that we didn't compute the eigenvalues 3 and -1 of A_1. Positive trace $3 - 1 = 2$, negative determinant $(3)(-1) = -3$. And $A_3 = -A_2$ is *negative* definite. The positive eigenvalues for A_2, two negative eigenvalues for A_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 is the trace so $a + c > 0$. Then a and c are both positive (if one of them is not positive, $ac - b^2 > 0$ will fail). Problem 1 reverses the reasoning to show that the tests 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 A is 3 by 3, $\det A > 0$ is the third part of the test. The next test requires *positive pivots*.

The eigenvalues of $A = A^{\mathrm{T}}$ ***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. The point is to recognize that ratio as the *second pivot* of A:

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} \xrightarrow[\text{The multiplier is } b/a]{\text{The first pivot is } a} \begin{bmatrix} a & b \\ 0 & c - \frac{b}{a}b \end{bmatrix} \qquad \textbf{\textit{The second pivot is}} \quad c - \frac{b^2}{a} = \frac{ac - b^2}{a}$$

This connects two big parts of linear algebra. ***Positive eigenvalues mean positive pivots and vice versa.*** We gave a proof for symmetric matrices of any size in the last section. 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.

$$A_1 = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 1 & -2 \\ -2 & 6 \end{bmatrix} \qquad A_3 = \begin{bmatrix} -1 & 2 \\ 2 & -6 \end{bmatrix}$$

pivots 1 and -3 **(indefinite)** — pivots 1 and 2 **(positive definite)** — pivots -1 and -2 **(negative definite)**

Here is a different way to look at symmetric matrices with positive eigenvalues.

Energy-based Definition

From $Ax = \lambda x$, multiply by x^T to get $x^T Ax = \lambda x^T x$. The right side is a positive λ times a positive number $x^T x = \|x\|^2$. So $x^T Ax$ is positive for any eigenvector.

The new idea is that $x^T Ax$ ***is positive for all nonzero vectors*** x, not just the eigenvectors. In many applications this number $x^T Ax$ (or $\frac{1}{2}x^T Ax$) 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 $x^T Ax > 0$.

Definition ***A is positive definite if $x^T Ax > 0$ for every nonzero vector x:***

$$x^T Ax = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2 > 0. \qquad (1)$$

The four entries a, b, b, c give the four parts of $x^T Ax$. 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 $x^T Ax$. This energy-based definition leads to a basic fact:

If A and B are symmetric positive definite, so is $A + B$.

Reason: $x^T(A + B)x$ is simply $x^T Ax + x^T Bx$. Those two terms are positive (for $x \neq \mathbf{0}$) so $A + B$ is also positive definite. The pivots and eigenvalues are not easy to follow when matrices are added, but the energies just add.

x^TAx also connects with our final way to recognize a positive definite matrix. Start with any matrix R, possibly rectangular. We know that $A = R^TR$ is square and symmetric. More than that, A will be positive definite when R has independent columns:

If the columns of R are independent, then $A = R^TR$ is positive definite.

Again eigenvalues and pivots are not easy. But the number x^TAx is the same as x^TR^TRx. That is exactly $(Rx)^T(Rx)$—another important proof by parenthesis! That vector Rx is not zero when $x \neq 0$ (this is the meaning of independent columns). Then x^TAx is the positive number $\|Rx\|^2$ and the matrix A 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 R^TR). Then come the applications.

When a symmetric matrix has one of these five properties, it has them all:

1. All ***n pivots*** are positive.
2. All ***n upper left determinants*** are positive.
3. All ***n eigenvalues*** are positive.
4. x^TAx is positive except at $x = 0$. This is the ***energy-based*** definition.
5. A equals R^TR for a matrix R with ***independent columns***.

The "upper left determinants" are 1 by 1, 2 by 2, ..., n by n. The last one is the determinant of the complete matrix A. This remarkable theorem ties together the whole linear algebra course—at least for symmetric matrices. We believe that two examples are more helpful than a detailed proof (we nearly have a proof already).

Example 1 Test these matrices A and B for positive definiteness:

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

Solution The pivots of A 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 A are $2-\sqrt{2}$ and 2 and $2+\sqrt{2}$, all positive. That completes tests **1**, **2**, and **3**.

We can write x^TAx as a sum of three squares. The pivots 2, $\frac{3}{2}$, $\frac{4}{3}$ appear outside the squares. The multipliers $-\frac{1}{2}$ and $-\frac{2}{3}$ from elimination are inside the squares:

$$x^TAx = 2\left(x_1^2 - x_1x_2 + x_2^2 - x_2x_3 + x_3^2\right) \qquad \textbf{Rewrite with squares}$$

$$= 2\left(x_1 - \tfrac{1}{2}x_2\right)^2 + \tfrac{3}{2}\left(x_2 - \tfrac{2}{3}x_3\right)^2 + \tfrac{4}{3}(x_3)^2. \qquad \textbf{This sum is positive.}$$

I have two candidates to suggest for R. Either one will show that $A = R^{\mathrm{T}}R$ is positive definite. R can be a rectangular first difference matrix, 4 by 3, to produce those second differences $-1, 2, -1$ in A:

$$\boldsymbol{A = R^{\mathrm{T}}R} \qquad \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

The three columns of this R are independent. A is positive definite.

Another R comes from $A = LDL^{\mathrm{T}}$ (the symmetric version of $A = LU$). 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 $\sqrt{D}$ with L.**

$$LDL^{\mathrm{T}} = \begin{bmatrix} 1 & & \\ -\frac{1}{2} & 1 & \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & \frac{3}{2} & \\ & & \frac{4}{3} \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & \\ & 1 & -\frac{2}{3} \\ & & 1 \end{bmatrix} = (L\sqrt{D})(L\sqrt{D})^{\mathrm{T}} = R^{\mathrm{T}}R. \quad (2)$$

R is the Cholesky factor

This choice of R has square roots (not so beautiful). But it is the only R that is 3 by 3 and upper triangular. It is the "Cholesky factor" of A and it is computed by MATLAB's command $R = \text{chol}(A)$. In applications, the rectangular R is how we build A and this Cholesky R is how we break it apart.

Eigenvalues give the symmetric choice $R = Q\sqrt{\Lambda}Q^{\mathrm{T}}$. This is also successful with $R^{\mathrm{T}}R = Q\Lambda Q^{\mathrm{T}} = A$. All these tests show that the $-1, 2, -1$ matrix A is positive definite.

Now turn to B, where 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 is still 3. The 3 by 3 determinant involves b:

$$\det B = 4 + 2b - 2b^2 = (1 + b)(4 - 2b) \quad \text{must be positive.}$$

At $b = -1$ and $b = 2$ we get $\det B = 0$. *Between $b = -1$ and $b = 2$ the matrix is positive definite.* The corner entry $b = 0$ in the first matrix A was safely between.

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 $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}0\boldsymbol{x} = 0$. These matrices on the edge are called *positive semidefinite*. Here are two examples (not invertible):

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \text{ are positive semidefinite.}$$

A has eigenvalues 5 and 0. Its upper left determinants are 1 and 0. Its rank is only 1. This matrix A factors into $R^{\mathrm{T}}R$ with **dependent columns** in R:

$$\begin{array}{l} \textbf{Dependent columns} \\ \textbf{Positive semidefinite} \end{array} \quad \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = R^{\mathrm{T}}R.$$

If 4 is increased by any small number, the matrix will become positive definite.

The cyclic B also has zero determinant (computed above when $b = -1$). It is singular. The eigenvector $\boldsymbol{x} = (1, 1, 1)$ has $B\boldsymbol{x} = \mathbf{0}$ and $\boldsymbol{x}^{\mathrm{T}}B\boldsymbol{x} = 0$. Vectors $\boldsymbol{x}$ in all other directions do give positive energy. This B can be written as $R^{\mathrm{T}}R$ in many ways, but R will always have *dependent* columns, with $(1, 1, 1)$ in its nullspace:

Second differences A
from first differences $R^{\mathrm{T}}R$
Cyclic A from cyclic R

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

Positive semidefinite matrices have all $\lambda \geq 0$ and all $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} \geq 0$. Those weak inequalities ($\geq$ instead of $>$) include positive definite matrices and the singular matrices at the edge.

First Application: The Ellipse $ax^2 + 2bxy + cy^2 = 1$

Think of a tilted ellipse $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = 1$. 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 $A = Q\Lambda Q^{-1} = Q\Lambda Q^{\mathrm{T}}$:

1. The tilted ellipse is associated with A. Its equation is $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = 1$.

2. The lined-up ellipse is associated with Λ. Its equation is $X^{\mathrm{T}}\Lambda X = 1$.

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:

$$\text{The equation is} \quad \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 1. \qquad \text{The matrix is} \quad A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}.$$

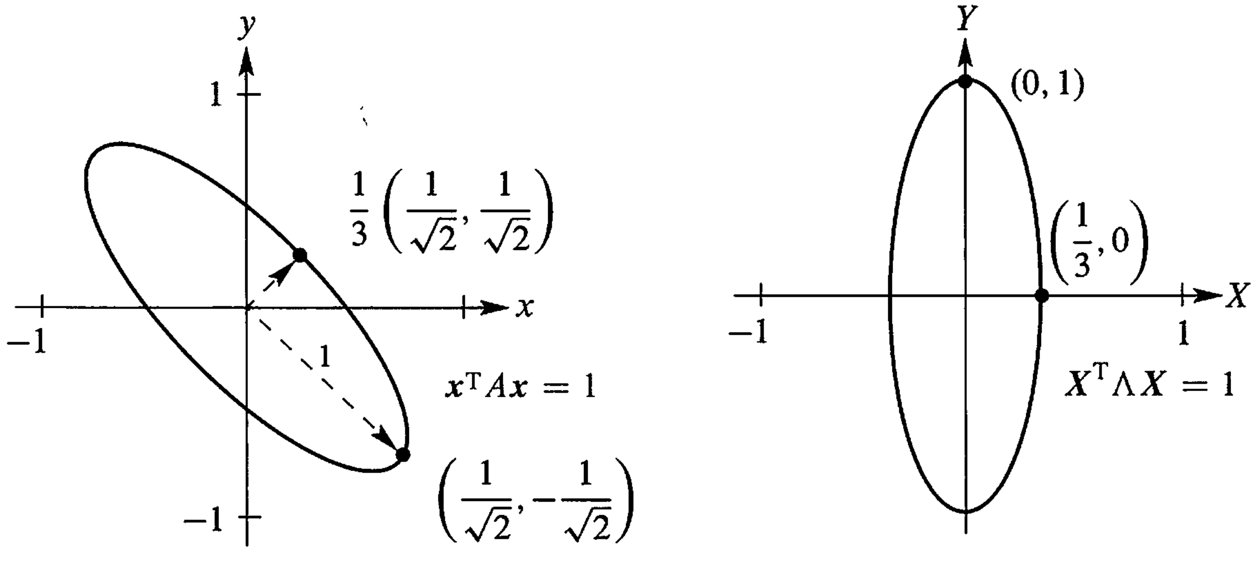


Figure 6.7: The tilted ellipse $5x^2 + 8xy + 5y^2 = 1$. Lined up it is $9X^2 + Y^2 = 1$.

The eigenvectors are $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\-1\end{bmatrix}$. Divide by $\sqrt{2}$ for unit vectors. Then $A = Q\Lambda Q^{\mathrm{T}}$:

Eigenvectors in Q
Eigenvalues 9 and 1
$$\begin{bmatrix}5 & 4\\4 & 5\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}\begin{bmatrix}\mathbf{9} & 0\\0 & \mathbf{1}\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}.$$

Now multiply by $\begin{bmatrix}x & y\end{bmatrix}$ on the left and $\begin{bmatrix}x\\y\end{bmatrix}$ on the right to get back to $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$:

$$x^{\mathrm{T}}Ax = \textbf{sum of squares} \quad 5x^2 + 8xy + 5y^2 = 9\left(\frac{x+y}{\sqrt{2}}\right)^2 + 1\left(\frac{x-y}{\sqrt{2}}\right)^2. \qquad (3)$$

The coefficients are not the pivots 5 and 9/5 from D, they are the eigenvalues 9 and 1 from Λ. Inside *these* squares are the eigenvectors $(1,1)/\sqrt{2}$ and $(1,-1)/\sqrt{2}$.

The axes of the tilted ellipse point along the eigenvectors. This explains why $A = Q\Lambda Q^{\mathrm{T}}$ 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:

$$\textbf{Lined up} \qquad \frac{x+y}{\sqrt{2}} = X \quad \text{and} \quad \frac{x-y}{\sqrt{2}} = Y \quad \text{and} \quad 9X^2 + Y^2 = 1.$$

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 λ_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 A. In the XY system, the axes are along the eigenvectors of Λ—the coordinate axes. All comes from $A = Q\Lambda Q^{\mathrm{T}}$.

Suppose $A = Q\Lambda Q^{\mathrm{T}}$ is positive definite, so $\lambda_i > 0$. The graph of $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = 1$ is an ellipse:

$$\begin{bmatrix}x & y\end{bmatrix} Q\Lambda Q^{\mathrm{T}}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}X & Y\end{bmatrix}\Lambda\begin{bmatrix}X\\Y\end{bmatrix} = \lambda_1 X^2 + \lambda_2 Y^2 = 1.$$

The axes point along eigenvectors. The half-lengths are $1/\sqrt{\lambda_1}$ and $1/\sqrt{\lambda_2}$.

$A = I$ gives the circle $x^2 + y^2 = 1$. If one eigenvalue is negative (exchange 4's and 5's in A), we don't have an ellipse. The sum of squares becomes a *difference of squares*: $9X^2 - Y^2 = 1$. This indefinite matrix gives a *hyperbola*. For a negative definite matrix like $A = -I$, with both λ's negative, the graph of $-x^2 - y^2 = 1$ has no points at all.

■ 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 $x^T Ax$ is then a "bowl" going up from $x = \mathbf{0}$:

$$x^T Ax = ax^2 + 2bxy + cy^2 \text{ is positive except at } (x, y) = (0, 0).$$

4. $A = R^T R$ is automatically positive definite if R has independent columns.
5. The ellipse $x^T Ax = 1$ has its axes along the eigenvectors of A. Lengths $1/\sqrt{\lambda}$.

■ WORKED EXAMPLES ■

6.5 A The great factorizations of a symmetric matrix are $A = LDL^T$ from pivots and multipliers, and $A = Q\Lambda Q^T$ from eigenvalues and eigenvectors. Show that $x^T Ax > 0$ for all nonzero x exactly when the pivots and eigenvalues are positive. Try these n by n tests on pascal(6) and ones(6) and hilb(6) and other matrices in MATLAB's gallery.

Solution To prove $x^T Ax > 0$, put parentheses into $x^T LDL^T x$ and $x^T Q\Lambda Q^T x$:

$$x^T Ax = (L^T x)^T D(L^T x) \qquad \text{and} \qquad x^T Ax = (Q^T x)^T \Lambda (Q^T x).$$

If x is nonzero, then $y = L^T x$ and $z = Q^T x$ are nonzero (those matrices are invertible). So $x^T Ax = y^T Dy = z^T \Lambda z$ becomes a sum of squares and A is shown as positive definite:

$$\begin{array}{llllllll} \textbf{Pivots} & x^T Ax & = & y^T Dy & = & d_1 y_1^2 + \cdots + d_n y_n^2 & > & 0 \\ \textbf{Eigenvalues} & x^T Ax & = & z^T \Lambda z & = & \lambda_1 z_1^2 + \cdots + \lambda_n z_n^2 & > & 0 \end{array}$$

MATLAB has a gallery of unusual matrices (type help gallery) and here are four:

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 two eigenvalues very near zero. **Hilbert matrix** $x^T Hx = \int_0^1 (x_1 + x_2 s + \cdots + x_6 s^5)^2 \, ds > 0$, $H_{ij} = 1/(i + j + 1)$.

rand(6)+rand(6)′ can be positive definite or not. Experiments gave only 2 in 20000.

$n = 20000$; $p = 0$; for $k = 1:n$, $A =$ rand(6); $p = p +$ all(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:

$$\begin{bmatrix} I & 0 \\ -B^T A^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & C - B^T A^{-1} B \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & S \end{bmatrix} \tag{4}$$

Those two blocks A and S must be positive definite. Their pivots are the pivots of M.

6.5 C ***Second 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 $(x,y) = (0,0)$?

Solution For $f(x)$, the test for a minimum comes from calculus: $df/dx = 0$ and $d^2f/dx^2 > 0$. Moving to two variables x and y produces a symmetric matrix H. It contains the four second derivatives of $F(x,y)$. Positive f'' changes to positive definite H:

Second derivative matrix $$H = \begin{bmatrix} \partial^2 F/\partial x^2 & \partial^2 F/\partial x \partial y \\ \partial^2 F/\partial y \partial x & \partial^2 F/\partial y^2 \end{bmatrix}$$

$F(x,y)$ has a minimum if H is positive definite. Reason: H reveals the important terms $ax^2 + 2bxy + cy^2$ near $(x,y) = (0,0)$. The second derivatives of F are $2a, 2b, 2b, 2c$!

6.5 D Find the eigenvalues of the $-1, 2, -1$ tridiagonal n by n matrix K (my favorite).

Solution The best way is to guess λ and $\boldsymbol{x}$. Then check $K\boldsymbol{x} = \lambda\boldsymbol{x}$. 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 K is like a *second derivative*, and those eigenvalues are much easier to see:

Eigenvalues $\lambda_1, \lambda_2, \ldots$
Eigenfunctions $y_1, y_2, \ldots$
$$\frac{d^2y}{dx^2} = \lambda y(x) \quad \textbf{with} \quad \begin{matrix} y(0) = 0 \\ y(1) = 0 \end{matrix} \tag{5}$$

Try $y = \sin cx$. Its second derivative is $y'' = -c^2 \sin cx$. So the eigenvalue will be $\lambda = -c^2$, provided $y(x)$ satisfies the end point conditions $y(0) = 0 = y(1)$.

Certainly $\sin 0 = 0$ (this is where cosines are eliminated by $\cos 0 = 1$). At $x = 1$, we need $y(1) = \sin c = 0$. The number c must be $k\pi$, a multiple of π, and λ is $-c^2$:

Eigenvalues $\lambda = -k^2\pi^2$
Eigenfunctions $y = \sin k\pi x$
$$\frac{d^2}{dx^2} \sin k\pi x = -k^2\pi^2 \sin k\pi x. \tag{6}$$

Now we go back to the matrix K and guess its eigenvectors. They come from $\sin k\pi x$ at n points $x = h, 2h, \ldots, nh$, equally spaced between 0 and 1. The spacing Δx is $h = 1/(n+1)$, so the $(n+1)$st point comes out at $(n+1)h = 1$. Multiply that sine vector s by K:

Eigenvector of K = sine vector s
$$\begin{matrix} Ks = \lambda s = (2 - 2\cos k\pi h)\, s \\ s = (\sin k\pi h, \ldots, \sin nk\pi h). \end{matrix} \tag{7}$$

I will leave that multiplication $Ks = \lambda s$ as a challenge problem. Notice what is important:

1. All eigenvalues $2 - 2\cos k\pi h$ are positive and K is positive definite.

2. The **sine matrix** S has orthogonal columns = eigenvectors $s_1, \ldots, s_n$ of K.

Discrete Sine Transform
The j, k entry is $\sin jk\pi h$

$$S = \begin{bmatrix} \sin \pi h & & \sin k\pi h & \\ \vdots & \cdots & \vdots & \cdots \\ \sin n\pi h & & \sin nk\pi h & \end{bmatrix}$$

Those eigenvectors are orthogonal just like the eigenfunctions: $\int_0^1 \sin j\pi x \sin k\pi x \, dx = 0$.

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 $c > b^2/a$ is also positive.

(i) λ_1 and λ_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 λ_1, λ_2.

2 Which of A_1, A_2, A_3, A_4 has two positive eigenvalues? Use the test, don't compute the λ's. Find an $\boldsymbol{x}$ so that $\boldsymbol{x}^{\mathrm{T}} A_1 \boldsymbol{x} < 0$, so A_1 fails the test.

$$A_1 = \begin{bmatrix} 5 & 6 \\ 6 & 7 \end{bmatrix} \quad A_2 = \begin{bmatrix} -1 & -2 \\ -2 & -5 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & 10 \\ 10 & 100 \end{bmatrix} \quad A_4 = \begin{bmatrix} 1 & 10 \\ 10 & 101 \end{bmatrix}.$$

3 For which numbers b and c are these matrices positive definite?

$$A = \begin{bmatrix} 1 & b \\ b & 9 \end{bmatrix} \quad A = \begin{bmatrix} 2 & 4 \\ 4 & c \end{bmatrix} \quad A = \begin{bmatrix} c & b \\ b & c \end{bmatrix}.$$

With the pivots in D and multiplier in L, factor each A into LDL^{T}.

4 What is the quadratic $f = ax^2 + 2bxy + cy^2$ for each of these matrices? Complete the square to write f as a sum of one or two squares $d_1(\quad)^2 + d_2(\quad)^2$.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 9 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}.$$

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. The minimum is not 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 A produces this f? What are its eigenvalues?

7 Test to see if $R^{\mathrm{T}}R$ is positive definite in each case:

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

8 The function $f(x,y) = 3(x+2y)^2 + 4y^2$ is positive except at $(0,0)$. What is the matrix in $f = [x \;\; y]A[x \;\; y]^{\mathrm{T}}$? Check that the pivots of A are 3 and 4.

9 Find the 3 by 3 matrix A and its pivots, rank, eigenvalues, and determinant:

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 4(x_1 - x_2 + 2x_3)^2.$$

10 Which 3 by 3 symmetric matrices A and B produce these quadratics?

$$x^{\mathrm{T}}Ax = 2(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_2x_3). \quad \text{Why is } A \text{ positive definite?}$$
$$x^{\mathrm{T}}Bx = 2(x_1^2 + x_2^2 + x_3^2 - x_1x_2 - x_1x_3 - x_2x_3). \quad \text{Why is } B \text{ semidefinite?}$$

11 Compute the three upper left determinants of A to establish positive definiteness. Verify that their ratios give the second and third pivots.

$$\textbf{Pivots = ratios of determinants} \qquad A = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 5 & 3 \\ 0 & 3 & 8 \end{bmatrix}.$$

12 For what numbers c and d are A and B positive definite? Test the 3 determinants:

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

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 A is positive definite then A^{-1} is positive definite.* Best proof: The eigenvalues of A^{-1} are positive because _____. *Second proof* (only for 2 by 2):

$$\text{The entries of } A^{-1} = \frac{1}{ac - b^2} \begin{bmatrix} c & -b \\ -b & a \end{bmatrix} \quad \text{pass the determinant tests} \quad \underline{\qquad}.$$

15 *If A and B are positive definite, their sum $A + B$ is positive definite.* Pivots and eigenvalues are not convenient for $A + B$. Better to prove $x^{\mathrm{T}}(A+B)x > 0$. Or if $A = R^{\mathrm{T}}R$ and $B = S^{\mathrm{T}}S$, show that $A + B = [R \;\; S]^{\mathrm{T}} \begin{bmatrix} R \\ S \end{bmatrix}$ with independent columns.

16 A positive definite matrix cannot have a zero (or even worse, a negative number) on its diagonal. Show that this matrix fails to have $x^{\mathrm{T}}Ax > 0$:

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 4 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ is not positive when } (x_1, x_2, x_3) = (\quad , \quad , \quad).$$

17 A diagonal entry a_{jj} of a symmetric matrix cannot be smaller than all the λ's. If it were, then $A - a_{jj}I$ would have _____ eigenvalues and would be positive definite. But $A - a_{jj}I$ has a _____ on the main diagonal.

18 If $Ax = \lambda x$ then $x^{\mathrm{T}}Ax =$ _____. If $x^{\mathrm{T}}Ax > 0$, prove that $\lambda > 0$.

19 Reverse Problem 18 to show that *if all* $\lambda > 0$ *then* $x^{\mathrm{T}}Ax > 0$. 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* $x_i^{\mathrm{T}}x_j = 0$. Then $x^{\mathrm{T}}Ax$ is

$$(c_1x_1 + \cdots + c_nx_n)^{\mathrm{T}}(c_1\lambda_1x_1 + \cdots + c_n\lambda_nx_n) = c_1^2\lambda_1x_1^{\mathrm{T}}x_1 + \cdots + c_n^2\lambda_nx_n^{\mathrm{T}}x_n > 0.$$

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 A and B have all $\lambda > 0$ (therefore positive definite)?

$$A = \begin{bmatrix} s & -4 & -4 \\ -4 & s & -4 \\ -4 & -4 & s \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} t & 3 & 0 \\ 3 & t & 4 \\ 0 & 4 & t \end{bmatrix}.$$

22 From $A = Q\Lambda Q^{\mathrm{T}}$ compute the positive definite symmetric square root $Q\Lambda^{1/2}Q^{\mathrm{T}}$ of each matrix. Check that this square root gives $R^2 = A$:

$$A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 10 & 6 \\ 6 & 10 \end{bmatrix}.$$

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 $\lambda_1x^2 + \lambda_2y^2 = 1$? The ellipse $9x^2 + 4y^2 = 1$ has axes with half-lengths $a =$ _____ and $b =$ _____.

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 A.

25 With positive pivots in D, the factorization $A = LDL^{\mathrm{T}}$ becomes $L\sqrt{D}\sqrt{D}L^{\mathrm{T}}$. (Square roots of the pivots give $D = \sqrt{D}\sqrt{D}$.) Then $C = \sqrt{D}L^{\mathrm{T}}$ yields the ***Cholesky factorization*** $A = C^{\mathrm{T}}C$ which is "symmetrized $L\,U$":

$$\text{From} \quad C = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{find } A. \qquad \text{From} \quad A = \begin{bmatrix} 4 & 8 \\ 8 & 25 \end{bmatrix} \quad \text{find } C = \mathbf{chol}(A).$$

26 In the Cholesky factorization $A = C^{\mathrm{T}}C$, with $C^{\mathrm{T}} = L\sqrt{D}$, the square roots of the pivots are on the diagonal of C. Find C (upper triangular) for

$$A = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 8 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 7 \end{bmatrix}.$$

27 The symmetric factorization $A = LDL^{\mathrm{T}}$ means that $\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}LDL^{\mathrm{T}}\boldsymbol{x}$:

$$\begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 1 & 0 \\ b/a & 1 \end{bmatrix}\begin{bmatrix} a & 0 \\ 0 & (ac-b^2)/a \end{bmatrix}\begin{bmatrix} 1 & b/a \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

The left side is $ax^2 + 2bxy + cy^2$. The right side is $a\left(x + \frac{b}{a}y\right)^2 + \underline{\qquad} y^2$. The second pivot completes the square! Test with $a = 2, b = 4, c = 10$.

28 Without multiplying $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$, find

(a) the determinant of A (b) the eigenvalues of A
(c) the eigenvectors of A (d) a reason why A 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 H_1 and H_2:

$$\textbf{Test for minimum}\quad H = \begin{bmatrix} \partial^2F/\partial x^2 & \partial^2F/\partial x\partial y \\ \partial^2F/\partial y\partial x & \partial^2F/\partial y^2 \end{bmatrix} \text{ is positive definite}$$

H_1 is positive definite so F_1 is concave up (= convex). Find the minimum point of F_1 and 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 $(0, 0)$?

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.

8.3 Applications of Markov Matrix---:)https://github.com/AvinashShrivastav/MATHEMATICS-FOR-COMPUTI

8.3 Markov Matrices, Population, and Economics

This section is about ***positive matrices***: every $a_{ij} > 0$. The key fact is quick to state: ***The largest eigenvalue is real and positive and so is its eigenvector.*** In economics and ecology and population dynamics and random walks, that fact leads a long way:

$$\textbf{Markov} \quad \lambda_{\max} = 1 \qquad \textbf{Population} \quad \lambda_{\max} > 1 \qquad \textbf{Consumption} \quad \lambda_{\max} < 1$$

$\lambda_{\max}$ controls the powers of A. We will see this first for $\lambda_{\max} = 1$.

Markov Matrices

Suppose we multiply a positive vector $\boldsymbol{u}_0 = (a, 1-a)$ again and again by this A:

Markov matrix $$A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \qquad \boldsymbol{u}_1 = A\boldsymbol{u}_0 \qquad \boldsymbol{u}_2 = A\boldsymbol{u}_1 = A^2\boldsymbol{u}_0$$

After k steps we have $A^k\boldsymbol{u}_0$. The vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \ldots$ will approach a "*steady state*" $\boldsymbol{u}_\infty = (.6, .4)$. This final outcome does not depend on the starting vector: ***For every $\boldsymbol{u}_0$ we converge to the same $\boldsymbol{u}_\infty$***. The question is why.

The steady state equation $A\boldsymbol{u}_\infty = \boldsymbol{u}_\infty$ makes $\boldsymbol{u}_\infty$ ***an eigenvector with eigenvalue* 1**:

Steady state $$\begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \begin{bmatrix} .6 \\ .4 \end{bmatrix} = \begin{bmatrix} .6 \\ .4 \end{bmatrix}.$$

Multiplying by A does not change $\boldsymbol{u}_\infty$. But this does not explain why all vectors $\boldsymbol{u}_0$ lead to $\boldsymbol{u}_\infty$. Other examples might have a steady state, but it is not necessarily attractive:

Not Markov $$B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{has the unattractive steady state} \quad B\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

In this case, the starting vector $\boldsymbol{u}_0 = (0, 1)$ will give $\boldsymbol{u}_1 = (0, 2)$ and $\boldsymbol{u}_2 = (0, 4)$. The second components are doubled. In the language of eigenvalues, B has $\lambda = 1$ but also $\lambda = 2$— this produces instability. The component of $\boldsymbol{u}$ along that unstable eigenvector is multiplied by λ, and $|\lambda| > 1$ means blowup.

This section is about two special properties of A that guarantee a stable steady state. These properties define a ***Markov matrix***, and A above is one particular example:

Markov matrix	**1.** ***Every entry of A is nonnegative.*** **2.** ***Every column of A adds to* 1.**

B did not have Property **2**. When A is a Markov matrix, two facts are immediate:

1. Multiplying a nonnegative $\boldsymbol{u}_0$ by A produces a nonnegative $\boldsymbol{u}_1 = A\boldsymbol{u}_0$.

2. If the components of $\boldsymbol{u}_0$ add to 1, so do the components of $\boldsymbol{u}_1 = A\boldsymbol{u}_0$.

Reason: The components of u_0 add to 1 when $\begin{bmatrix}1 & \cdots & 1\end{bmatrix}u_0 = 1$. This is true for each column of A by Property 2. Then by matrix multiplication $[1 \ \ldots \ 1]A = [1 \ \ldots \ 1]$:

$$\textbf{Components of } Au_0 \textbf{ add to 1} \qquad [1 \ \cdots \ 1]Au_0 = [1 \ \cdots \ 1]u_0 = 1.$$

The same facts apply to $u_2 = Au_1$ and $u_3 = Au_2$. ***Every vector $A^k u_0$ is nonnegative with components adding to*** 1. These are "***probability vectors.***" The limit u_∞ is also a probability vector—but we have to prove that there is a limit. We will show that $\lambda_{\max} = 1$ for a positive Markov matrix.

Example 1 The fraction of rental cars in Denver starts at $\frac{1}{50} = .02$. The fraction outside Denver is .98. Every month, 80% of the Denver cars stay in Denver (and 20% leave). Also 5% of the outside cars come in (95% stay outside). This means that the fractions $u_0 = (.02, .98)$ are multiplied by A:

$$\textbf{First month} \qquad A = \begin{bmatrix} .80 & .05 \\ .20 & .95 \end{bmatrix} \quad \text{leads to} \quad u_1 = Au_0 = A\begin{bmatrix} .02 \\ .98 \end{bmatrix} = \begin{bmatrix} .065 \\ .935 \end{bmatrix}.$$

Notice that $.065 + .935 = 1$. All cars are accounted for. Each step multiplies by A:

$$\textbf{Next month} \qquad u_2 = Au_1 = (.09875,\ .90125). \text{ This is } A^2u_0.$$

All these vectors are positive because A is positive. Each vector u_k will have its components adding to 1. The first component has grown from .02 and cars are moving toward Denver. What happens in the long run?

This section involves powers of matrices. The understanding of A^k was our first and best application of diagonalization. Where A^k can be complicated, the diagonal matrix Λ^k is simple. The eigenvector matrix S connects them: A^k equals $S\Lambda^k S^{-1}$. The new application to Markov matrices uses the eigenvalues (in Λ) and the eigenvectors (in S). We will show that u_∞ is an eigenvector corresponding to $\lambda = 1$.

Since every column of A adds to 1, nothing is lost or gained. We are moving rental cars or populations, and no cars or people suddenly appear (or disappear). The fractions add to 1 and the matrix A keeps them that way. The question is how they are distributed after k time periods—which leads us to A^k.

Solution $A^k u_0$ gives the fractions in and out of Denver after k steps. We diagonalize A to understand A^k. The eigenvalues are $\lambda = 1$ and .75 (the trace is 1.75).

$$Ax = \lambda x \qquad A\begin{bmatrix} .2 \\ .8 \end{bmatrix} = 1\begin{bmatrix} .2 \\ .8 \end{bmatrix} \quad \textbf{and} \quad A\begin{bmatrix} -1 \\ 1 \end{bmatrix} = .75\begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

The starting vector u_0 combines x_1 and x_2, in this case with coefficients 1 and .18:

$$\textbf{Combination of eigenvectors} \qquad u_0 = \begin{bmatrix} .02 \\ .98 \end{bmatrix} = \begin{bmatrix} .2 \\ .8 \end{bmatrix} + .18\begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Now multiply by A to find u_1. The eigenvectors are multiplied by $\lambda_1 = 1$ and $\lambda_2 = .75$:

$$\textbf{Each } x \textbf{ is multiplied by } \lambda \qquad u_1 = 1\begin{bmatrix} .2 \\ .8 \end{bmatrix} + (.75)(.18)\begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Every month, another .75 multiplies the vector x_2. The eigenvector x_1 is unchanged:

After k steps $$u_k = A^k u_0 = \begin{bmatrix} .2 \\ .8 \end{bmatrix} + (.75)^k(.18)\begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

This equation reveals what happens. ***The eigenvector x_1 with $\lambda = 1$ is the steady state.*** The other eigenvector x_2 disappears because $|\lambda| < 1$. The more steps we take, the closer we come to $u_\infty = (.2, .8)$. In the limit, $\frac{2}{10}$ of the cars are in Denver and $\frac{8}{10}$ are outside. This is the pattern for Markov chains, even starting from $u_0 = (0, 1)$:

If A is a *positive* Markov matrix (entries $a_{ij} > 0$, each column adds to **1**), then $\lambda_1 = 1$ is larger than any other eigenvalue. The eigenvector x_1 is the ***steady state***:

$$u_k = x_1 + c_2(\lambda_2)^k x_2 + \cdots + c_n(\lambda_n)^k x_n \quad \textit{always approaches} \quad u_\infty = x_1.$$

The first point is to see that $\lambda = 1$ is an eigenvalue of A. *Reason*: Every column of $A - I$ adds to $1 - 1 = 0$. The rows of $A - I$ add up to the zero row. Those rows are linearly dependent, so $A - I$ is singular. Its determinant is zero and $\lambda = 1$ is an eigenvalue.

The second point is that no eigenvalue can have $|\lambda| > 1$. With such an eigenvalue, the powers A^k would grow. But A^k is also a Markov matrix! A^k has nonnegative entries still adding to 1—and that leaves no room to get large.

A lot of attention is paid to the possibility that another eigenvalue has $|\lambda| = 1$.

Example 2 $A = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ has no steady state because $\lambda_2 = -1$.

This matrix sends all cars from inside Denver to outside, and vice versa. The powers A^k alternate between A and I. The second eigenvector $x_2 = (-1, 1)$ will be multiplied by $\lambda_2 = -1$ at every step—and does not become smaller: No steady state.

Suppose the entries of A or any power of A are all *positive*—zero is not allowed. In this "regular" or "primitive" case, $\lambda = 1$ is strictly larger than any other eigenvalue. The powers A^k approach the rank one matrix that has the steady state in every column.

Example 3 (**"Everybody moves"**) Start with three groups. At each time step, half of group 1 goes to group 2 and the other half goes to group 3. The other groups also *split in half and move*. Take one step from the starting populations p_1, p_2, p_3:

New populations $$u_1 = Au_0 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}p_2 + \frac{1}{2}p_3 \\ \frac{1}{2}p_1 + \frac{1}{2}p_3 \\ \frac{1}{2}p_1 + \frac{1}{2}p_2 \end{bmatrix}.$$

A is a Markov matrix. Nobody is born or lost. A contains zeros, which gave trouble in Example 2. But after two steps in this new example, the zeros disappear from A^2:

Two-step matrix $$u_2 = A^2 u_0 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}.$$

The eigenvalues of A are $\lambda_1 = 1$ (because A is Markov) and $\lambda_2 = \lambda_3 = -\frac{1}{2}$. For $\lambda = 1$, ***the eigenvector*** $\boldsymbol{x}_1 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ ***will be the steady state.*** When three equal populations split in half and move, the populations are again equal. Starting from $\boldsymbol{u}_0 = (8, 16, 32)$, the Markov chain approaches its steady state:

$$\boldsymbol{u}_0 = \begin{bmatrix} 8 \\ 16 \\ 32 \end{bmatrix} \quad \boldsymbol{u}_1 = \begin{bmatrix} 24 \\ 20 \\ 12 \end{bmatrix} \quad \boldsymbol{u}_2 = \begin{bmatrix} 16 \\ 18 \\ 22 \end{bmatrix} \quad \boldsymbol{u}_3 = \begin{bmatrix} 20 \\ 19 \\ 17 \end{bmatrix}.$$

The step to $\boldsymbol{u}_4$ will split some people in half. This cannot be helped. The total population is $8 + 16 + 32 = 56$ at every step. The steady state is 56 times $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. You can see the three populations approaching, but never reaching, their final limits $56/3$.

Challenge Problem 6.7.16 created a Markov matrix A from the number of links between websites. The steady state $\boldsymbol{u}$ will give the Google rankings. ***Google finds*** $\boldsymbol{u}_\infty$ ***by a random walk that follows links*** (*random surfing*). That eigenvector comes from counting the fraction of visits to each website—a quick way to compute the steady state.

The size $|\lambda_2|$ of the next largest eigenvalue controls the speed of convergence to steady state.

Perron-Frobenius Theorem

One matrix theorem dominates this subject. The Perron-Frobenius Theorem applies when all $a_{ij} \geq 0$. There is no requirement that columns add to 1. We prove the neatest form, when all $a_{ij} > 0$.

Perron-Frobenius for $A > 0$ ***All numbers in*** $A\boldsymbol{x} = \lambda_{\max}\boldsymbol{x}$ ***are strictly positive.***

Proof The key idea is to look at all numbers t such that $A\boldsymbol{x} \geq t\boldsymbol{x}$ for some nonnegative vector $\boldsymbol{x}$ (other than $\boldsymbol{x} = \boldsymbol{0}$). We are allowing inequality in $A\boldsymbol{x} \geq t\boldsymbol{x}$ in order to have many positive candidates t. For the largest value $t_{\max}$ (which is attained), we will show that ***equality holds***: $A\boldsymbol{x} = t_{\max}\boldsymbol{x}$.

Otherwise, if $A\boldsymbol{x} \geq t_{\max}\boldsymbol{x}$ is not an equality, multiply by A. Because A is positive that produces a strict inequality $A^2\boldsymbol{x} > t_{\max}A\boldsymbol{x}$. Therefore the positive vector $\boldsymbol{y} = A\boldsymbol{x}$ satisfies $A\boldsymbol{y} > t_{\max}\boldsymbol{y}$, and $t_{\max}$ could be increased. This contradiction forces the equality $A\boldsymbol{x} = t_{\max}\boldsymbol{x}$, and we have an eigenvalue. Its eigenvector $\boldsymbol{x}$ is positive because on the left side of that equality, $A\boldsymbol{x}$ is sure to be positive.

To see that no eigenvalue can be larger than $t_{\max}$, suppose $A\boldsymbol{z} = \lambda\boldsymbol{z}$. Since λ and $\boldsymbol{z}$ may involve negative or complex numbers, we take absolute values: $|\lambda||\boldsymbol{z}| = |A\boldsymbol{z}| \leq A|\boldsymbol{z}|$ by the "triangle inequality." This $|\boldsymbol{z}|$ is a nonnegative vector, so $|\lambda|$ is one of the possible candidates t. Therefore $|\lambda|$ cannot exceed $t_{\max}$—which must be $\lambda_{\max}$.

Population Growth

Divide the population into three age groups: age < 20, age 20 to 39, and age 40 to 59. At year T the sizes of those groups are n_1, n_2, n_3. Twenty years later, the sizes have changed for two reasons:

1. **Reproduction** $n_1^{\text{new}} = F_1 n_1 + F_2 n_2 + F_3 n_3$ gives a new generation

2. **Survival** $n_2^{\text{new}} = P_1 n_1$ and $n_3^{\text{new}} = P_2 n_2$ gives the older generations

The fertility rates are F_1, F_2, F_3 (F_2 largest). The *Leslie matrix* A might look like this:

$$\begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}^{\textbf{new}} = \begin{bmatrix} F_1 & F_2 & F_3 \\ P_1 & 0 & 0 \\ 0 & P_2 & 0 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \begin{bmatrix} .04 & \mathbf{1.1} & .01 \\ .98 & 0 & 0 \\ 0 & \mathbf{.92} & 0 \end{bmatrix} \begin{bmatrix} n_1 \\ \boldsymbol{n_2} \\ n_3 \end{bmatrix}.$$

This is population projection in its simplest form, the same matrix A at every step. In a realistic model, A will change with time (from the environment or internal factors). Professors may want to include a fourth group, age ≥ 60, but we don't allow it.

The matrix has $A \geq 0$ but not $A > 0$. The Perron-Frobenius theorem still applies because $A^3 > 0$. The largest eigenvalue is $\lambda_{\max} \approx 1.06$. You can watch the generations move, starting from $n_2 = 1$ in the middle generation:

$$\textbf{eig}(A) = \begin{matrix} \mathbf{1.06} \\ -1.01 \\ -0.01 \end{matrix} \qquad A^2 = \begin{bmatrix} 1.08 & \mathbf{0.05} & .00 \\ 0.04 & \mathbf{1.08} & .01 \\ 0.90 & 0 & 0 \end{bmatrix} \qquad A^3 = \begin{bmatrix} 0.10 & \mathbf{1.19} & .01 \\ 0.06 & \mathbf{0.05} & .00 \\ 0.04 & \mathbf{0.99} & .01 \end{bmatrix}.$$

A fast start would come from $\boldsymbol{u}_0 = (0, 1, 0)$. That middle group will reproduce 1.1 and also survive .92. The newest and oldest generations are in $\boldsymbol{u}_1 = (1.1, 0, .92) =$ column 2 of A. Then $\boldsymbol{u}_2 = A\boldsymbol{u}_1 = A^2\boldsymbol{u}_0$ is the second column of A^2. The early numbers (transients) depend a lot on $\boldsymbol{u}_0$, but ***the asymptotic growth rate*** $\lambda_{\max}$ ***is the same from every start.*** Its eigenvector $\boldsymbol{x} = (.63, .58, .51)$ shows all three groups growing steadily together.

Caswell's book on *Matrix Population Models* emphasizes sensitivity analysis. The model is never exactly right. If the F's or P's in the matrix change by 10%, does $\lambda_{\max}$ go below 1 (which means extinction)? Problem 19 will show that a matrix change ΔA produces an eigenvalue change $\Delta\lambda = \boldsymbol{y}^{\mathrm{T}}(\Delta A)\boldsymbol{x}$. Here $\boldsymbol{x}$ and $\boldsymbol{y}^{\mathrm{T}}$ are the right and left eigenvectors of A. So $\boldsymbol{x}$ is a column of S and $\boldsymbol{y}^{\mathrm{T}}$ is a row of S^{-1}.

Linear Algebra in Economics: The Consumption Matrix

A long essay about linear algebra in economics would be out of place here. A short note about one matrix seems reasonable. The ***consumption matrix*** tells how much of each input goes into a unit of output. This describes the manufacturing side of the economy.

Consumption matrix We have n industries like chemicals, food, and oil. To produce a unit of chemicals may require .2 units of chemicals, .3 units of food, and .4 units of oil. Those numbers go into row 1 of the consumption matrix A:

$$\begin{bmatrix} \text{chemical output} \\ \text{food output} \\ \text{oil output} \end{bmatrix} = \begin{bmatrix} .2 & .3 & .4 \\ .4 & .4 & .1 \\ .5 & .1 & .3 \end{bmatrix} \begin{bmatrix} \text{chemical input} \\ \text{food input} \\ \text{oil input} \end{bmatrix}.$$

Row 2 shows the inputs to produce food—a heavy use of chemicals and food, not so much oil. Row 3 of A shows the inputs consumed to refine a unit of oil. The real consumption matrix for the United States in 1958 contained 83 industries. The models in the 1990's are much larger and more precise. We chose a consumption matrix that has a convenient eigenvector.

Now comes the question: Can this economy meet demands y_1, y_2, y_3 for chemicals, food, and oil? To do that, the inputs p_1, p_2, p_3 will have to be higher—because part of $\boldsymbol{p}$ is consumed in producing $\boldsymbol{y}$. The input is $\boldsymbol{p}$ and the consumption is $A\boldsymbol{p}$, which leaves the output $\boldsymbol{p} - A\boldsymbol{p}$. This net production is what meets the demand $\boldsymbol{y}$:

Problem Find a vector $\boldsymbol{p}$ such that $\quad \boldsymbol{p} - A\boldsymbol{p} = \boldsymbol{y} \quad$ or $\quad \boldsymbol{p} = (I - A)^{-1}\boldsymbol{y}$.

Apparently the linear algebra question is whether $I - A$ is invertible. But there is more to the problem. The demand vector $\boldsymbol{y}$ is nonnegative, and so is A. *The production levels in* $\boldsymbol{p} = (I - A)^{-1}\boldsymbol{y}$ *must also be nonnegative*. The real question is:

When is $(I - A)^{-1}$ a nonnegative matrix?

This is the test on $(I - A)^{-1}$ for a productive economy, which can meet any positive demand. If A is small compared to I, then $A\boldsymbol{p}$ is small compared to $\boldsymbol{p}$. There is plenty of output. If A is too large, then production consumes more than it yields. In this case the external demand $\boldsymbol{y}$ cannot be met.

"Small" or "large" is decided by the largest eigenvalue λ_1 of A (which is positive):

If $\lambda_1 > 1$ then $(I - A)^{-1}$ has negative entries
If $\lambda_1 = 1$ then $(I - A)^{-1}$ fails to exist
If $\lambda_1 < 1$ then $(I - A)^{-1}$ is nonnegative as desired.

The main point is that last one. The reasoning uses a nice formula for $(I - A)^{-1}$, which we give now. The most important infinite series in mathematics is the ***geometric series*** $1 + x + x^2 + \cdots$. This series adds up to $1/(1 - x)$ provided x lies between -1 and 1. When $x = 1$ the series is $1 + 1 + 1 + \cdots = \infty$. When $|x| \ge 1$ the terms x^n don't go to zero and the series has no chance to converge.

The nice formula for $(I - A)^{-1}$ is the ***geometric series of matrices***:

Geometric series $\quad (I - A)^{-1} = I + A + A^2 + A^3 + \cdots.$

If you multiply the series $S = I + A + A^2 + \cdots$ by A, you get the same series except for I. Therefore $S - AS = I$, which is $(I - A)S = I$. The series adds to $S = (I - A)^{-1}$ if it converges. ***And it converges if all eigenvalues of A have $|\lambda| < 1$.***

In our case $A \geq 0$. All terms of the series are nonnegative. Its sum is $(I - A)^{-1} \geq 0$.

Example 4 $A = \begin{bmatrix} .2 & .3 & .4 \\ .4 & .4 & .1 \\ .5 & .1 & .3 \end{bmatrix}$ has $\lambda_{\max} = .9$ and $(I - A)^{-1} = \frac{1}{93}\begin{bmatrix} 41 & 25 & 27 \\ 33 & 36 & 24 \\ 34 & 23 & 36 \end{bmatrix}$.

This economy is productive. A is small compared to I, because $\lambda_{\max}$ is .9. To meet the demand $\boldsymbol{y}$, start from $\boldsymbol{p} = (I - A)^{-1}\boldsymbol{y}$. Then $A\boldsymbol{p}$ is consumed in production, leaving $\boldsymbol{p} - A\boldsymbol{p}$. This is $(I - A)\boldsymbol{p} = \boldsymbol{y}$, and the demand is met.

Example 5 $A = \begin{bmatrix} 0 & 4 \\ 1 & 0 \end{bmatrix}$ has $\lambda_{\max} = 2$ and $(I - A)^{-1} = -\frac{1}{3}\begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}$.

This consumption matrix A is too large. Demands can't be met, because production consumes more than it yields. The series $I + A + A^2 + \ldots$ does not converge to $(I - A)^{-1}$ because $\lambda_{\max} > 1$. The series is growing while $(I - A)^{-1}$ is actually negative.

In the same way $1 + 2 + 4 + \cdots$ is not really $1/(1 - 2) = -1$. But not entirely false!

Problem Set 8.3

Questions 1–12 are about Markov matrices and their eigenvalues and powers.

1 Find the eigenvalues of this Markov matrix (their sum is the trace):

$$A = \begin{bmatrix} .90 & .15 \\ .10 & .85 \end{bmatrix}.$$

What is the steady state eigenvector for the eigenvalue $\lambda_1 = 1$?

2 Diagonalize the Markov matrix in Problem 1 to $A = S\Lambda S^{-1}$ by finding its other eigenvector:

$$A = \begin{bmatrix} & \\ & \end{bmatrix}\begin{bmatrix} 1 & \\ & .75 \end{bmatrix}\begin{bmatrix} & \\ & \end{bmatrix}.$$

What is the limit of $A^k = S\Lambda^k S^{-1}$ when $\Lambda^k = \begin{bmatrix} 1 & 0 \\ 0 & .75^k \end{bmatrix}$ approaches $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$?

3 What are the eigenvalues and steady state eigenvectors for these Markov matrices?

$$A = \begin{bmatrix} 1 & .2 \\ 0 & .8 \end{bmatrix} \quad A = \begin{bmatrix} .2 & 1 \\ .8 & 0 \end{bmatrix} \quad A = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}.$$

4 For every 4 by 4 Markov matrix, what eigenvector of A^{T} corresponds to the (known) eigenvalue $\lambda = 1$?

5 Every year 2% of young people become old and 3% of old people become dead. (No births.) Find the steady state for

$$\begin{bmatrix} \text{young} \\ \text{old} \\ \text{dead} \end{bmatrix}_{k+1} = \begin{bmatrix} .98 & .00 & 0 \\ .02 & .97 & 0 \\ .00 & .03 & 1 \end{bmatrix} \begin{bmatrix} \text{young} \\ \text{old} \\ \text{dead} \end{bmatrix}_k .$$

6 For a Markov matrix, the sum of the components of $\boldsymbol{x}$ equals the sum of the components of $A\boldsymbol{x}$. If $A\boldsymbol{x} = \lambda \boldsymbol{x}$ with $\lambda \neq 1$, prove that the components of this non-steady eigenvector $\boldsymbol{x}$ add to zero.

7 Find the eigenvalues and eigenvectors of A. Explain why A^k approaches A^∞:

$$A = \begin{bmatrix} .8 & .3 \\ .2 & .7 \end{bmatrix} \qquad A^\infty = \begin{bmatrix} .6 & .6 \\ .4 & .4 \end{bmatrix}.$$

Challenge problem: Which Markov matrices produce that steady state $(.6, .4)$?

8 The steady state eigenvector of a permutation matrix is $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. This is *not* approached when $\boldsymbol{u}_0 = (0, 0, 0, 1)$. What are $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$ and $\boldsymbol{u}_4$? What are the four eigenvalues of P, which solve $\lambda^4 = 1$?

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

9 Prove that the square of a Markov matrix is also a Markov matrix.

10 If $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ is a Markov matrix, its eigenvalues are 1 and ______. The steady state eigenvector is $\boldsymbol{x}_1 =$ ______.

11 Complete A to a Markov matrix and find the steady state eigenvector. When A is a symmetric Markov matrix, why is $\boldsymbol{x}_1 = (1, \ldots, 1)$ its steady state?

$$A = \begin{bmatrix} .7 & .1 & .2 \\ .1 & .6 & .3 \\ \_ & \_ & \_ \end{bmatrix}.$$

12 A Markov differential equation is not $d\boldsymbol{u}/dt = A\boldsymbol{u}$ but $d\boldsymbol{u}/dt = (A - I)\boldsymbol{u}$. The diagonal is negative, the rest of $A - I$ is positive. The columns add to zero.

Find the eigenvalues of $B = A - I = \begin{bmatrix} -.2 & .3 \\ .2 & -.3 \end{bmatrix}$. Why does $A - I$ have $\lambda = 0$?

When $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$ multiply $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, what is the steady state as $t \to \infty$?