## Introduction to Diagonalization

Suppose $T: \mathbb{R}^{n} \rightarrow \mathbb{R}^{n}$ is a linear transformation, and that $T(\boldsymbol{x})=D \boldsymbol{x}$ where $D$ is an $n \times n$ diagonal matrix. Such transformations are particularly easy to understand, and can be very useful in applications.

The effect of multiplication by $D$ is just a rescaling in the directions of the basis coordinate vectors $e_{1}, e_{2}, \ldots, e_{n}$ :
$\begin{aligned} T\left(\boldsymbol{e}_{i}\right)=D \boldsymbol{e}_{\boldsymbol{i}} & =\left[\begin{array}{ccccc}\lambda_{1} & 0 & 0 & 0 & 0 \\ 0 & \lambda_{2} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \lambda_{n}\end{array}\right]\left[\begin{array}{c}0 \\ \vdots \\ \vdots \\ 1 \\ \vdots \\ 0\end{array}\right] \swarrow \text { row } i \\ & =1\left(\operatorname{col}_{i} D\right)=1 \cdot\left[\begin{array}{c}0 \\ \vdots \\ \vdots \\ \lambda_{i} \\ \vdots \\ 0\end{array}\right]=\lambda_{i}\left[\begin{array}{c}0 \\ \vdots \\ \vdots \\ 1 \\ \vdots \\ 0\end{array}\right] \swarrow \text { row } i, \quad \text { so } T\left(\boldsymbol{e}_{\boldsymbol{i}}\right)=D \boldsymbol{e}_{\boldsymbol{i}}=\lambda_{i} \boldsymbol{e}_{\boldsymbol{i}}\end{aligned}$
just rescales each basis vector: $e_{i}$ maps to a multiple of itself, and the rescaling factor $=$ the entry $\lambda_{i}$ from the diagonal of $D$.

Since the $\boldsymbol{e}_{i}$ 's are a basis, each vector $\boldsymbol{x}=\left[\begin{array}{c}x_{1} \\ x_{2} \\ \vdots \\ \vdots \\ x_{n}\end{array}\right]=x_{1} \boldsymbol{e}_{\mathbf{1}}+\ldots+x_{n} \boldsymbol{e}_{n}$, so

$$
T(\boldsymbol{x})=D \boldsymbol{x} \quad=D\left(x_{1} \boldsymbol{e}_{1}+\ldots+x_{n} \boldsymbol{e}_{n}\right)=x_{1} D\left(\boldsymbol{e}_{1}\right)+\ldots+x_{n} D\left(\boldsymbol{e}_{n}\right)
$$

$$
=\lambda_{1} x_{1} \boldsymbol{e}_{\mathbf{1}}+\ldots+\lambda_{n} x_{n} \boldsymbol{e}_{\boldsymbol{n}}=\left[\begin{array}{c}
\lambda_{1} x_{1} \\
\lambda_{2} x_{2} \\
\vdots \\
\vdots \\
\lambda_{n} x_{n}
\end{array}\right] .
$$

In other words the effect of $T(\boldsymbol{x})=D \boldsymbol{x}$ is just to rescale each coordinate $x_{i}$ of $\boldsymbol{x}$ by the corresponding factor $\lambda_{i}$.

Example: Suppose $T: \mathbb{R}^{2} \rightarrow \mathbb{R}^{2}$ is given by $T(\boldsymbol{x})=D \boldsymbol{x}=\left[\begin{array}{ll}3 & 0 \\ 0 & 6\end{array}\right] \boldsymbol{x}$.
Then $\quad D \mathbf{e}_{1}=\left[\begin{array}{ll}3 & 0 \\ 0 & 6\end{array}\right]\left[\begin{array}{l}1 \\ 0\end{array}\right]=\left[\begin{array}{l}3 \\ 0\end{array}\right]=3 \boldsymbol{e}_{1}, D \mathbf{e}_{2}=\left[\begin{array}{ll}3 & 0 \\ 0 & 6\end{array}\right]\left[\begin{array}{l}0 \\ 1\end{array}\right]=\left[\begin{array}{l}0 \\ 6\end{array}\right]=6 \boldsymbol{e}_{2}$,
and $\quad D \boldsymbol{x}=\left[\begin{array}{ll}3 & 0 \\ 0 & 6\end{array}\right] \boldsymbol{x}=\left[\begin{array}{ll}3 & 0 \\ 0 & 6\end{array}\right]\left[\begin{array}{l}x_{1} \\ x_{2}\end{array}\right]=\left[\begin{array}{l}3 x_{1} \\ 6 x_{2}\end{array}\right]$.
$D$ simply rescales each coordinate of $\boldsymbol{x}$ - by a factor of 3 in the direction of $\boldsymbol{e}_{1}$ (the horizontal direction) and by a factor of 6 in the direction of $e_{2}$ (the vertical direction). In this particular example, the diagonal entries of $D$ are both bigger than 1 , so multiplying $x$ by $D$ "stretches" the coordinates of $x$ in each of the coordinate directions, $e_{1}$ and $e_{2}$.

You can see this in the figure below where $\left[\begin{array}{l}1 \\ 1\end{array}\right] \mapsto D\left[\begin{array}{l}1 \\ 1\end{array}\right]=\left[\begin{array}{l}3 \\ 6\end{array}\right]$ :


Example (the example is for $\mathbb{R}^{2}$, but the same idea applies just as well for $\mathbb{R}^{n}$ ). Suppose $T(\boldsymbol{x})=D \boldsymbol{x}=\left[\begin{array}{cc}1 & 0 \\ 0 & \frac{1}{2}\end{array}\right] \boldsymbol{x}$. Consider the difference equation

$$
x_{k+1}=D x_{k}
$$

It's easy to see the long-term behavior of this system since multiplication by $D$ just rescales coordinates: if we start with $\quad x_{0}=\left[\begin{array}{l}a \\ b\end{array}\right]$, then

$$
\boldsymbol{x}_{1}=D\left[\begin{array}{l}
a \\
b
\end{array}\right]=\left[\begin{array}{l}
1 \cdot a \\
\frac{1}{2} \cdot b
\end{array}\right], \quad \boldsymbol{x}_{2}=D \boldsymbol{x}_{1}=D\left[\begin{array}{c}
1 \cdot a \\
\frac{1}{2} \cdot b
\end{array}\right]=\left[\begin{array}{c}
1^{2} \cdot a \\
\left(\frac{1}{2}\right)^{2} \cdot b
\end{array}\right]
$$

$$
\begin{aligned}
& \boldsymbol{x}_{3}=D \boldsymbol{x}_{2}=D\left[\begin{array}{c}
1^{2} \cdot a \\
\left(\frac{1}{2}\right)^{2} \cdot b
\end{array}\right]=\left[\begin{array}{c}
1^{3} \cdot a \\
\left(\frac{1}{2}\right)^{3} \cdot b
\end{array}\right] \\
& \boldsymbol{x}_{\boldsymbol{k}}=\left[\begin{array}{c}
1^{k} \cdot a \\
\left(\frac{1}{2}\right)^{k} \cdot b
\end{array}\right]=\left[\begin{array}{c}
a \\
\left(\frac{1}{2}\right)^{k} \cdot b
\end{array}\right]
\end{aligned}
$$

Since $\left(\frac{1}{2}\right)^{k} \rightarrow 0$ as $k \rightarrow \infty$, we can immediately see the long term behavior of the system:

$$
\boldsymbol{x}_{k} \rightarrow\left[\begin{array}{l}
a \\
0
\end{array}\right] \text { as } k \rightarrow \infty
$$

In the same way, for any diagonal matrix $D=\left[\begin{array}{cc}\lambda_{1} & 0 \\ 0 & \lambda_{2}\end{array}\right]$ and $x_{0}=\left[\begin{array}{l}a \\ b\end{array}\right]$, we get

$$
\boldsymbol{x}_{\boldsymbol{k}}=\left[\begin{array}{c}
\lambda_{1}^{k} a \\
\lambda_{2}^{k} b
\end{array}\right], \text { so what happens to } \boldsymbol{x}_{\boldsymbol{k}} \text { is determined by the size of }
$$

$\lambda_{1}, \lambda_{2}$ : for example, if both $\left|\lambda_{i}\right|<1$. then $\boldsymbol{x}_{\boldsymbol{k}} \rightarrow \mathbf{0}$. The long term behavior of such a system can be very important in applications, and it depends on the diagonal entries of $D$.

Suppose $T: \mathbb{R}^{n} \rightarrow \mathbb{R}^{n}$, where $T(\boldsymbol{x})=A \boldsymbol{x}$. If $A$ is not diagonal, then sometimes it's possible to find a new basis $\mathcal{B}$ so that a change to working in $\mathcal{B}$ coordinates "diagonalizes" $A$. Roughly, this means that
$\left(^{*}\right)\left\{\begin{array}{lll}\boldsymbol{x} & \mapsto A \boldsymbol{x} \quad \text { (in standard coordinates) is the same as } \\ {[\boldsymbol{x}]_{\mathcal{B}}} & \mapsto D[\boldsymbol{x}]_{\mathcal{B}} & \text { (expressed instead in } \mathcal{B} \text { coordinates) }\end{array}\right.$
To make this more precise:
For a basis $\mathcal{B}$, recall that $P_{\mathcal{B}}$ is the matrix whose columns are the new basis vectors in $\mathcal{B}$. Then standard coordinates and $\mathcal{B}$ coordinates are related by the equations

$$
P_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}=\boldsymbol{x} \quad \text { and } \quad[\boldsymbol{x}]_{\mathcal{B}}=P_{\mathcal{B}}^{-1} \boldsymbol{x}
$$

$\underline{\text { Sometimes we can find a basis } \mathcal{B}}$ and a diagonal matrix $D$ so that if we:

1) first convert $\boldsymbol{x}$ to $\mathcal{B}$ coordinates
2) then multiply by $D$ (rescale $\mathcal{B}$ coordinates)
3) then convert back to standard coordinates

$$
\begin{array}{r}
P_{\mathcal{B}}^{-1} \boldsymbol{x} \\
D P_{\mathcal{B}}^{-1} \boldsymbol{x} \\
P_{\mathcal{B}} D P_{\mathcal{B}}^{-1} \boldsymbol{x}
\end{array}
$$

we get the same result as $A \boldsymbol{x}$. In other words, $A \boldsymbol{x}=P_{\mathcal{B}} D P_{\mathcal{B}}^{-1} \boldsymbol{x}$
In that case, the effect of the mapping $\boldsymbol{x} \mapsto A \boldsymbol{x}$ is a rescaling of the $\mathcal{B}$ coordinates of $\boldsymbol{x}$ with rescaling factors $=$ the entries on the diagonal of $D$.

This leads to the official definition:

Suppose $A$ is an $n \times n$ matrix. $A$ is called diagonalizable if there is an invertible $n \times n$ matrix $P$ and a diagonal matrix $D$ such that

$$
A=P D P^{-1}
$$

(When this happens, the columns of $P$ can be used as a new basis $\mathcal{B}$ for $\mathbb{R}^{n}$ and then $P=P_{\mathcal{B}}=$ the "change of coordinates matrix" so that things look just as they did in the preceding paragraphs.)

It is not always possible to diagonalize an $n \times n$ matrix $A$. We will see more about this in Chapter 5. But the examples below will lead us to one statement about when it is possible.

Example Suppose $T(\boldsymbol{x})=A \boldsymbol{x}$, where $A=\left[\begin{array}{ll}4 & 2 \\ 1 & 5\end{array}\right]$. It's a fact (just check by multiplying, if you like) that

$$
\begin{aligned}
& A=\left[\begin{array}{ll}
4 & 2 \\
1 & 5
\end{array}\right]=\left[\begin{array}{rr}
2 & 1 \\
-1 & 1
\end{array}\right]\left[\begin{array}{ll}
3 & 0 \\
0 & 6
\end{array}\right]\left[\begin{array}{rr}
\frac{1}{3} & -\frac{1}{3} \\
\frac{1}{3} & \frac{2}{3}
\end{array}\right] \\
& =\begin{array}{ccc}
\uparrow & \uparrow & \uparrow \\
P & D & P^{-1}
\end{array}
\end{aligned}
$$

(for now, don't worry about where the matrix $P$ came from. Such questions come up in Chapter 5)
$P$ is invertible, so its columns are linearly independent and span $\mathbb{R}^{2}$ and we can use the columns as a new basis for $\mathbb{R}^{2}: \mathcal{B}=\left\{\boldsymbol{b}_{1}, \boldsymbol{b}_{2}\right\}=\left\{\left[\begin{array}{r}2 \\ -1\end{array}\right],\left[\begin{array}{l}1 \\ 1\end{array}\right]\right\}$. Then $P$ is the matrix that the textbook calls $P_{\mathcal{B}}=$ "the change of coordinates matrix from $\mathcal{B}$ coordinates to standard coordinates" :

$$
P_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}=\boldsymbol{x}
$$

This new basis lets us see more clearly how $A$ operates. Suppose $\boldsymbol{x}=\left[\begin{array}{l}2 \\ 1\end{array}\right]$. Then

$$
A\left[\begin{array}{l}
2 \\
1
\end{array}\right]=\left[\begin{array}{ll}
4 & 2 \\
1 & 5
\end{array}\right]\left[\begin{array}{l}
2 \\
1
\end{array}\right]=\left[\begin{array}{c}
10 \\
7
\end{array}\right]
$$

But we can also calculate $A \boldsymbol{x}$ in a more roundabout way - but one which gives new insight:

$$
\begin{aligned}
A\left[\begin{array}{l}
2 \\
1
\end{array}\right]= & {\left[\begin{array}{ll}
4 & 2 \\
1 & 5
\end{array}\right]\left[\begin{array}{l}
2 \\
1
\end{array}\right] }
\end{aligned}=\left[\begin{array}{rr}
2 & 1 \\
-1 & 1
\end{array}\right]\left[\begin{array}{ll}
3 & 0 \\
0 & 6
\end{array}\right]\left[\begin{array}{rr}
\frac{1}{3} & -\frac{1}{3} \\
\frac{1}{3} & \frac{2}{3}
\end{array}\right]\left[\begin{array}{l}
2 \\
1
\end{array}\right]=\left[\begin{array}{c}
10 \\
7
\end{array}\right]
$$

In the second calculation, look at what happens to $\boldsymbol{x}$, step-by-step, as each matrix, in turn, does its work:

$$
\begin{aligned}
& \boldsymbol{x}=\left[\begin{array}{l}
2 \\
1
\end{array}\right] \mapsto\left[\begin{array}{rr}
\frac{1}{3} & -\frac{1}{3} \\
\frac{1}{3} & \frac{2}{3}
\end{array}\right]\left[\begin{array}{l}
2 \\
1
\end{array}\right]=\left[\begin{array}{l}
\frac{1}{3} \\
\frac{4}{3}
\end{array}\right] \cdot \underline{\text { Then }}\left[\begin{array}{ll}
3 & 0 \\
0 & 6
\end{array}\right]\left[\begin{array}{l}
\frac{1}{3} \\
\frac{4}{3}
\end{array}\right]=\left[\begin{array}{l}
1 \\
8
\end{array}\right] \\
& \text { multiplication by } \uparrow \quad \uparrow \quad \uparrow \\
& P_{B}^{-1} \text { converts } \quad[\boldsymbol{x}]_{\mathcal{B}} \text { the diagonal matrix stretched } \\
& \text { standard coordinates } \quad D \text { stretches the new } \quad \mathcal{B} \text {-coordinates } \\
& \text { into } \mathcal{B} \text {-coordinates } \mathcal{B} \text {-coordinates } \\
& \text { by factors of } 3 \text { and } 6 \\
& \text { (in the coordinate directions } \\
& \text { corresponding to } \boldsymbol{b}_{1} \text { and } \boldsymbol{b}_{2} \text { ) } \\
& \text { Finally, } \quad\left[\begin{array}{rr}
2 & 1 \\
-1 & 1
\end{array}\right]\left[\begin{array}{l}
1 \\
8
\end{array}\right]=\left[\begin{array}{c}
10 \\
7
\end{array}\right] \\
& \text { multiplication by } P_{B} \\
& \text { converts the (stretched) } \\
& \mathcal{B} \text {-coordinates back into } \\
& \text { standard coordinates }
\end{aligned}
$$

Use the figure on the following page to check each of the preceding steps graphically, as accurately as you can.

1) $\boldsymbol{x}=\left[\begin{array}{l}2 \\ 1\end{array}\right]$ has $\mathcal{B}$-coordinates $[\boldsymbol{x}]_{\mathcal{B}}=\left[\begin{array}{l}\frac{1}{3} \\ \frac{4}{3}\end{array}\right]$
2) Stretching $[\boldsymbol{x}]_{\mathcal{B}}$ by factors of 3 and 6 in the $\boldsymbol{b}_{1}$ and $\boldsymbol{b}_{\boldsymbol{2}}$ directions gives the point with "stretched" $\mathcal{B}$-coordinates $\left[\begin{array}{l}1 \\ 8\end{array}\right]$.
3) Converting these $\mathcal{B}$-coordinates $\left[\begin{array}{l}1 \\ 8\end{array}\right]$ back to standard coordinates gives $\left[\begin{array}{c}10 \\ 7\end{array}\right]=A\left[\begin{array}{l}2 \\ 1\end{array}\right]$.


So the effect of $A$ on a point $\boldsymbol{x}$ (in standard coordinates) is the same as the effect of the diagonal matrix $D$ on the same point $\boldsymbol{x}$ (when described in $\mathcal{B}$-coordinates) : a rescaling of the $\mathcal{B}$ coordinates by a factor of 3 (in the direction of the "positive $\boldsymbol{b}_{1}$-axis") and by a factor of 6 (in the direction of the "positive $\boldsymbol{b}_{2}$-axis)."

Notice that $\quad \boldsymbol{b}_{1}=\left[\begin{array}{r}2 \\ -1\end{array}\right] \quad$ has $\mathcal{B}$ coordinates $\left[\begin{array}{l}1 \\ 0\end{array}\right]$ and $\boldsymbol{b}_{2}$ has $\mathcal{B}$ coordinates $\left[\begin{array}{l}0 \\ 1\end{array}\right]$
so (by the underlined remark in the preceding paragraph)
$A\left[\begin{array}{r}2 \\ -1\end{array}\right]=A \boldsymbol{b}_{\mathbf{1}}=$ the point with $\mathcal{B}$ coords. $\left[\begin{array}{l}3 \\ 0\end{array}\right]=3 \boldsymbol{b}_{\mathbf{1}}=3\left[\begin{array}{r}2 \\ -1\end{array}\right]$ (std coords.)
and

$$
A\left[\begin{array}{l}
1 \\
1
\end{array}\right]=A \boldsymbol{b}_{2}=\text { the point with } \mathcal{B} \text { coords. }\left[\begin{array}{l}
0 \\
6
\end{array}\right]=6 \boldsymbol{b}_{2}=6\left[\begin{array}{l}
1 \\
1
\end{array}\right](\text { in std coords })
$$

Again, it might be helpful to look at what's going on in more detail, using the same 3-step process. See the figure on the next page.

For example, for $\boldsymbol{b}_{\mathbf{1}}$ (follow each step using the preceding figure) :

$$
\begin{gathered}
A \boldsymbol{b}_{\mathbf{1}}=\left[\begin{array}{ll}
4 & 2 \\
1 & 5
\end{array}\right]\left[\begin{array}{r}
2 \\
-1
\end{array}\right]=\left[\begin{array}{cc}
2 & 1 \\
-1 & 1
\end{array}\right]\left[\begin{array}{ll}
3 & 0 \\
0 & 6
\end{array}\right]\left[\begin{array}{cc}
\frac{1}{3} & -\frac{1}{3} \\
\frac{1}{3} & \frac{2}{3}
\end{array}\right]\left[\begin{array}{c}
2 \\
-1
\end{array}\right] \\
=\left[\begin{array}{rr}
2 & 1 \\
-1 & 1
\end{array}\right]\left[\begin{array}{ll}
3 & 0 \\
0 & 6
\end{array}\right]\left[\begin{array}{l}
1 \\
0
\end{array}\right]=\left[\begin{array}{ll}
2 & 1 \\
-1 & 1
\end{array}\right]\left[\begin{array}{l}
3 \\
0
\end{array}\right]=\left[\begin{array}{c}
6 \\
-3
\end{array}\right]=3\left[\begin{array}{c}
2 \\
-1
\end{array}\right]=3 \boldsymbol{b}_{1} \\
\mathcal{B} \text {-coordinates } \\
\text { of } \boldsymbol{b}_{\mathbf{1}}=
\end{gathered} \begin{gathered}
\text { stretched } \\
\mathcal{B} \text {-coordinates }
\end{gathered} \begin{gathered}
\text { back into standard } \\
\text { coordinates again }
\end{gathered}
$$

"unit vector" in the $b_{1}$ direction
Thus the matrix $A$ maps each vector $b_{1}$ and $b_{2}$ to a scalar multiple of itself. Such (nonzero) vectors are called eigenvectors of $A$; the scalar 3 is called the eigenvalue associated with the eigenvector $\boldsymbol{b}_{1}$, and 6 is called the eigenvalue associated with the eigenvector $\boldsymbol{b}_{2}$.

The reason we were able to do such a nice analysis of how this matrix $A$ works is that we were able to write down a new basis $\mathcal{B}$ for $\mathbb{R}^{2}$ - a basis whose members are eigenvectors of $A$. (Remember, in the beginning, I simply gave you the matrix $P$ whose columns $\boldsymbol{b}_{1} \boldsymbol{b}_{2}$ turned out to be eigenvectors; we haven't discussed yet how you might start, from "", to find such a)

Here is the general definition.
Definition A nonzero vector $\boldsymbol{x}$ is called an eigenvector of the $n \times n$ matrix $A$ if $A \boldsymbol{x}=\lambda \boldsymbol{x}$ for some scalar $\lambda$. The scalar $\lambda$ is called an eigenvalue of $A$ (associated with the eigenvector $\boldsymbol{x}$ ). Eigenvalues and eigenvectors of $A$ are also called eigenvectors or eigenvalues of the transformation $T$ if $T(\boldsymbol{x})=A \boldsymbol{x}$,

It's traditional, in almost all books, to use the Greek letter $\lambda$ ("lambda") to denote an eigenvalue.

Eigenvalue and eigenvector are words with German roots. Some books call "eigenvectors" and "eigenvalues" by the loose translations "characteristic vectors" and "characteristic values.")

In the preceding $2 \times 2$ example there is nothing special about the specific eigenvalues 3 and 6 . They might be any other numbers $\lambda_{1}$ and $\lambda_{2}$ (where perhaps even $\lambda_{1}=\lambda_{2}$ ). To recap, in general:

If $A$ can be factored as $A=P\left[\begin{array}{cc}\lambda_{1} & 0 \\ 0 & \lambda_{2}\end{array}\right] P^{-1}$, where $P$ has columns $\boldsymbol{b}_{\mathbf{1}}, \boldsymbol{b}_{2}$, then $\boldsymbol{b}_{1}$ and $\boldsymbol{b}_{2}$ are eigenvectors of $A$ with eigenvalues $\lambda_{1}$ and $\lambda_{2}$ because:

We can choose $\mathcal{B}=\left\{\boldsymbol{b}_{\mathbf{1}}, \boldsymbol{b}_{2}\right\}$ as a new basis (why?) for $\mathbb{R}^{2}$, and then calculate

$$
A \boldsymbol{b}_{1}=P_{\mathcal{B}}\left[\begin{array}{cc}
\lambda_{1} & 0 \\
0 & \lambda_{2}
\end{array}\right] P_{\mathcal{B}}^{-1} \boldsymbol{b}_{1}=P_{\mathcal{B}}\left[\begin{array}{cc}
\lambda_{1} & 0 \\
0 & \lambda_{2}
\end{array}\right]\left[\begin{array}{l}
1 \\
0
\end{array}\right]
$$

$\mathcal{B}$-coordinates of $\boldsymbol{b}_{1}$

$$
\begin{aligned}
= & P_{\mathcal{B}}\left[\begin{array}{c}
\lambda_{1} \\
0
\end{array}\right]=\left[\begin{array}{ll}
\boldsymbol{b}_{1} & \boldsymbol{b}_{2}
\end{array}\right]\left[\begin{array}{c}
\lambda_{1} \\
0
\end{array}\right]=\lambda_{1} \boldsymbol{b}_{1}+0 \boldsymbol{b}_{2}=\lambda_{1} \boldsymbol{b}_{1} \\
& \mathcal{B} \text {-coordinates }
\end{aligned}
$$

Similarly, $A \boldsymbol{b}_{\mathbf{2}}=\lambda_{2} \boldsymbol{b}_{\mathbf{2}}$.
In this case, the geometric interpretation of how $A$ operates on a point $x$ is just as it was earlier: there is a rescaling of $\mathcal{B}$ coordinates, with rescaling factors $=\lambda_{1}$ and $\lambda_{2}$. Of course, if (say) $0<\lambda_{1}<1$ then multiplication by $D$ is a "contraction" of the first $\mathcal{B}$ coordinate of $\boldsymbol{x}$ (rather than a"stretch"); and if (say) $\lambda_{2}<0$, multiplication by $D$ reverses the sign of the second $\mathcal{B}$ coordinate of $\boldsymbol{x}$ (as well as either stretching or contracting).

The preceding discussion proves the following theorem (where $n=2$ ), and the discussion is exactly the same for $n \times n$ matrices (except that there are more matrix entries and vectors to write down, making the discuss look as if it were more complicated - it isn't !) So we have indicated the proof of the following theorem:

Theorem 1 Let $A$ be an $n \times n$ matrix
Suppose $A$ is diagonalizable, that is, suppose $A$ can be factored as

$$
A=P\left[\begin{array}{ccccc}
\lambda_{1} & 0 & 0 & 0 & 0 \\
0 & \lambda_{2} & 0 & \cdots & 0 \\
0 & \cdots & \lambda_{3} & \cdots & 0 \\
0 & \vdots & \vdots & \vdots & 0 \\
0 & 0 & 0 & 0 & \lambda_{n}
\end{array}\right] P^{-1}
$$

Then $\mathbb{R}^{n}$ has a basis that consists of eigenvectors of $A$.
This basis is $\mathcal{B}=\left\{\boldsymbol{b}_{\mathbf{1}}, \boldsymbol{b}_{\boldsymbol{2}}, \ldots \boldsymbol{b}_{\boldsymbol{n}}\right\}$, where the $\boldsymbol{b}_{\boldsymbol{i}}$ 's are the columns of $P$. These $\boldsymbol{b}_{\boldsymbol{i}}$ 's are eigenvectors of $A$, and their eigenvalues are the scalars of the diagonal of $D$ (in same order: $\lambda_{1}$ for $\boldsymbol{b}_{1}$, etc.).

It turns out that the converse of Theorem 1 is also true: if we start with a basis for $\mathbb{R}^{n}$ that consists of eigenvectors of $A$, then $A$ must be diagonalizable - that is, $A$ can be factored as $A=P D P^{-1}$.

The next example illustrates why this works.
Example Suppose $A=\left[\begin{array}{rr}7 & 2 \\ -4 & 1\end{array}\right]$. Consider that basis $\mathcal{B}=\left\{\left[\begin{array}{r}1 \\ -1\end{array}\right]\right.$ and $\left.\left[\begin{array}{r}1 \\ -2\end{array}\right]\right\}$.
The vectors $\boldsymbol{b}_{\mathbf{1}}$ and $\boldsymbol{b}_{\mathbf{2}}$ are eigenvectors for $A$ as we can easily check:

$$
A\left[\begin{array}{r}
1 \\
-1
\end{array}\right]=\left[\begin{array}{rr}
7 & 2 \\
-4 & 1
\end{array}\right]\left[\begin{array}{r}
1 \\
-1
\end{array}\right]=\left[\begin{array}{r}
5 \\
-5
\end{array}\right]=5\left[\begin{array}{r}
1 \\
-1
\end{array}\right]
$$

$$
\begin{gathered}
\text { so }\left[\begin{array}{r}
1 \\
-1
\end{array}\right] \text { is an eigenvector with eigenvalue } 5 \text {. Similarly, } \\
A\left[\begin{array}{r}
1 \\
-2
\end{array}\right]=\left[\begin{array}{rr}
7 & 2 \\
-4 & 1
\end{array}\right]\left[\begin{array}{r}
1 \\
-2
\end{array}\right]=\left[\begin{array}{r}
3 \\
-6
\end{array}\right]=3\left[\begin{array}{r}
1 \\
-2
\end{array}\right] \\
\text { so }\left[\begin{array}{r}
1 \\
-2
\end{array}\right] \text { is an eigenvector with eigenvalue } 3 .
\end{gathered}
$$

So $\mathcal{B}=\left\{\boldsymbol{b}_{\mathbf{1}}, \boldsymbol{b}_{2}\right\}$ is a basis for $\mathbb{R}^{2}$ consisting of eigenvectors of $A$.
Create a matrix $P=\left[\begin{array}{ll}\boldsymbol{b}_{1} & \boldsymbol{b}_{\mathbf{2}}\end{array}\right]=\left[\begin{array}{rr}1 & 1 \\ -1 & -2\end{array}\right]$. Because the columns of $P$ are linearly independent, $P$ is invertible.

Create $D=\left[\begin{array}{ll}5 & 0 \\ 0 & 3\end{array}\right]$, where the diagonal entries are the eigenvalues of $\boldsymbol{b}_{\mathbf{1}}, \boldsymbol{b}_{2}$ (written down in the same order: not $\left[\begin{array}{ll}3 & 0 \\ 0 & 5\end{array}\right]!!$ )

Then $A=P D P^{-1}$. To see that this factorization works, the easiest thing to do (rather than compute $P^{-1}$ ) is to multiply both sides of the proposed equality on the right by $P$ and check whether the equivalent equation $A P=P D$ is true:

$$
\begin{array}{ll}
A P=\left[\begin{array}{rr}
7 & 2 \\
-4 & 1
\end{array}\right]\left[\begin{array}{rr}
1 & 1 \\
-1 & -2
\end{array}\right] & =\left[\begin{array}{rr}
5 & 3 \\
-5 & -6
\end{array}\right] \text { and } \\
P D=\left[\begin{array}{rr}
1 & 1 \\
-1 & -2
\end{array}\right]\left[\begin{array}{ll}
5 & 0 \\
0 & 3
\end{array}\right] & =\left[\begin{array}{rr}
5 & 3 \\
-5 & -6
\end{array}\right]
\end{array}
$$

so the factorization is correct.
The general statement of the converse for Theorem 1 is
Theorem 2 Let $A$ be an $n \times n$ matrix.
Suppose $\mathbb{R}^{n}$ has a basis $\left\{\boldsymbol{b}_{\mathbf{1}}, \boldsymbol{b}_{\mathbf{2}}, \ldots, \boldsymbol{b}_{\boldsymbol{n}}\right\}$ that consists of eigenvectors of $A$.
Then $A$ is diagonalizable, that is $A=P D P^{-1}$ for some diagonal matrix $D$.
The matrix $P$ has the eigenvectors $\boldsymbol{b}_{1}, \boldsymbol{b}_{2}, \ldots, \boldsymbol{b}_{n}$ as its columns, and the diagonal matrix $D$ has on its diagonal the corresponding eigenvalues $\lambda_{1}, \lambda_{2}, \ldots, \lambda_{n}$ (written in the same order: $\lambda_{1}$ for $\boldsymbol{b}_{1}$, etc.)

Why is Theorem 2 true? The argument is given here for the case $n=2$, but it works in exactly the same way for $\mathbb{R}^{n}$. Not surprisingly, the argument in general looks almost exactly like the preceding example.

Suppose $A$ is $2 \times 2$ and we know that $\mathcal{B}=\left\{\boldsymbol{b}_{\mathbf{1}}, \boldsymbol{b}_{2}\right\}$ is a basis $\mathbb{R}^{2}$ consisting of eigenvectors of $A$. Let their corresponding eigenvalues be $\lambda_{1}$ and $\lambda_{2}$.

Create matrix $P=\left[\begin{array}{ll}\boldsymbol{b}_{1} & \boldsymbol{b}_{2}\end{array}\right]$ and $D=\left[\begin{array}{cc}\lambda_{1} & 0 \\ 0 & \lambda_{2}\end{array}\right]$. Since $\boldsymbol{b}_{\mathbf{1}}$ and $\boldsymbol{b}_{\mathbf{2}}$ are linearly independent, $P$ is invertible. To complete the proof that $A$ is diagonalizable, we that these matrices work: that $A=P D P^{-1}$.

As in the preceding example, we just need to verify that $A P=P D$, and, to do this, we simply need to remember the definition of matrix multiplication.

$$
\begin{aligned}
& P D= P\left[\begin{array}{cc}
\lambda_{1} & 0 \\
0 & \lambda_{2}
\end{array}\right]=\left[P\left[\begin{array}{c}
\lambda_{1} \\
0
\end{array}\right] P\left[\begin{array}{c}
0 \\
\lambda_{2}
\end{array}\right]\right] \\
&= {\left.\left[\begin{array}{ll}
\boldsymbol{b}_{1} & \boldsymbol{b}_{2}
\end{array}\right] \cdot\left[\begin{array}{c}
\lambda_{1} \\
0
\end{array}\right] \quad\left[\begin{array}{cc}
\boldsymbol{b}_{1} & \boldsymbol{b}_{2}
\end{array}\right] \cdot\left[\begin{array}{c}
0 \\
\lambda_{2}
\end{array}\right]\right]=\left[\begin{array}{ll}
\lambda_{1} \boldsymbol{b}_{\mathbf{1}} & \lambda_{2} \boldsymbol{b}_{2}
\end{array}\right], \text { and } } \\
& A P= A\left[\boldsymbol{b}_{\mathbf{1}} \boldsymbol{b}_{\mathbf{2}}\right]= \\
& {\left[A \boldsymbol{b}_{1} A \boldsymbol{b}_{2}\right]=\left[\begin{array}{lll}
\lambda_{1} \boldsymbol{b}_{1} & \lambda_{2} \boldsymbol{b}_{2}
\end{array}\right] } \\
& \xrightarrow{\text { because }} \boldsymbol{b}_{1} \text { is an eigenvector of } A \text { with eigenvalue } \lambda_{1}, \\
& \boldsymbol{b}_{2} \text { is an eigenvector of } A \text { with eigenvalue } \lambda_{2} .
\end{aligned}
$$

Therefore $A P$ and $P D$ have the same columns, so $A P=P D$

Question Based on this discussion, can you give an example of a transformation $T: \mathbb{R}^{2} \rightarrow \mathbb{R}^{2}$ where $T(\boldsymbol{x})=A \boldsymbol{x}$ and $A$ is not diagonalizable?

## The Big Picture, so far

In Chapter 4, we have been discussing vector spaces $V$ (where $V$ might not be $\mathbb{R}^{n}$ ). After discussing linear independence and spanning in this more general setting, we were led to the idea of a basis $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{n}\right\}$ for $V$. From that, the Unique Representation Theorem (p. 246) led to the idea of using $\mathcal{B}$ to create coordinates for each $\boldsymbol{x}$ in $V$. The coordinate vector $[\boldsymbol{x}]_{\mathcal{B}}$ is a vector in $\mathbb{R}^{n}$.

We saw that the mapping $\boldsymbol{x} \mapsto[\boldsymbol{x}]_{\mathcal{B}}$ is an isomorphism (a one-to-one, onto linear mapping) from the vector space $V$ to $\mathbb{R}^{n}$. This association preserves all vector space operations and all linear dependency relations. For example, a linearly independent set of vectors in $V$ has a linearly independent set of coordinate vectors in $\mathbb{R}^{n}$, and vice-versa.

In the special case when $V=\mathbb{R}^{n}$, we can might choose a basis $\mathcal{B}=\left\{\boldsymbol{b}_{1}, \ldots, \boldsymbol{b}_{n}\right\}$ different from the standard basis. Then a vector $\boldsymbol{x}$ in $\mathbb{R}^{n}$ gets new "nonstandard" coordinates $[\boldsymbol{x}]_{\mathcal{B}}$ relative to the basis $\mathcal{B}$. The matrix $P_{\mathcal{B}}=\left[\boldsymbol{b}_{1} \ldots \boldsymbol{b}_{\boldsymbol{n}}\right]$ is the "operator" that changes $\mathcal{B}$ coordinates into standard coordinates according to the formula $P_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}=\boldsymbol{x}$.

Since the columns of $P_{\mathcal{B}}$ are linearly independent, the change of coordinates matrix $P_{\mathcal{B}}$ is always invertible and $P_{\mathcal{B}}^{-1}$ is the "operator" that converts standard coordinates into $\mathcal{B}$-coordinates: $[\boldsymbol{x}]_{\mathcal{B}}=P_{\mathcal{B}}^{-1} \boldsymbol{x}$.

If a $2 \times 2$ matrix $A$ can be factored into the form $P D P^{-1}$, where $D$ is a diagonal matrix, then $A$ is called diagonalizable - because in that case $A$ "acts like a diagonal matrix" when computations are done relative to the basis $\mathcal{B}$ that consists of the columns. This led into the idea of eigenvectors and eigenvalues of the matrix $A$.

We proved that a $2 \times 2$ matrix $A$ is diagonalizable if and only if $\mathbb{R}^{2}$ has a basis consisting of eigenvectors of $A$ - and indicated that a completely similar proof works for an $n \times n$ matrix $A$ operating on $\mathbb{R}^{n}$.

The conceptual idea of diagonalization and its relation to a basis of eigenvectors is nicely motivated geometrically and not very hard. You may have noticed, however, that in the preceding examples:
a factorization of a given matrix $A$ into $P D P^{-1}$ was given, or
the eigenvalues and eigenvectors of $A$ were given so that $P$ and $D$ could be created.
But if you are simply given an $n \times n$ matrix $A$, then trying to find its eigenvectors and eigenvalues (if it has any at all!), and determining whether $\mathbb{R}^{n}$ does or does not have a basis consisting of eigenvectors of $A$ are harder questions. Be aware of those questions, but try to keep worries about them suppressed until we get into Chapter 5. For now just focus on the concept of diagonalization, what it means, and how diagonalization is connected to eigenvectors and eigenvalues.

