There are situations where we are guaranteed a basis of $\mathbb{R}^n$ made out eigenvectors of $A$:

**Theorem 1**: Let $A$ be an $n \times n$ matrix with distinct real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be corresponding (non-zero) eigenvectors, $A \mathbf{v}_j = \lambda_j \mathbf{v}_j$. Then the set

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$$

is linearly independent, and so is a basis for $\mathbb{R}^n$....this is one we can prove!

**Nifty proof**: Assume the vectors in the set are dependent. (we’ll end up with a contradiction)

$$\{\mathbf{v}_i\} \text{ independent}$$

$$\{\mathbf{v}_1, \mathbf{v}_2\} \text{ ?}$$

$$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$$

$$\vdots$$

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\} \text{ dependent}$$

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ be the first dependent set in this list.

Since it’s the first (shortest) set, $\mathbf{v}_p$ is a combination of the earlier $\mathbf{v}_j$’s

(1)

$$\mathbf{v}_p = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \ldots + c_p \mathbf{v}_p$$

not all $c_j = 0$

$$\Rightarrow A \mathbf{v}_p = A(c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \ldots + c_p \mathbf{v}_p)$$

$$\lambda_p \mathbf{v}_p = c_1 \lambda_1 \mathbf{v}_1 + c_2 \lambda_2 \mathbf{v}_2 + \ldots + c_p \lambda_p \mathbf{v}_p$$

(2)

$$\lambda_p \mathbf{v}_p = c_1 \lambda_1 \mathbf{v}_1 + c_2 \lambda_2 \mathbf{v}_2 + \ldots + c_p \lambda_p \mathbf{v}_p$$

From (2), $\lambda_p \neq 0$, if it was,

$$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\} \text{ is dependent}$$

(3)

$$\mathbf{v}_p = c_1 \frac{\lambda_1 \mathbf{v}_1}{\lambda_p} + c_2 \frac{\lambda_2 \mathbf{v}_2}{\lambda_p} + \ldots + c_{p-1} \frac{\lambda_{p-1} \mathbf{v}_{p-1}}{\lambda_p}$$

Eqn 1 - Eqn 3:

$$\mathbf{0} = c_1 \left(1 - \frac{\lambda_1}{\lambda_p}\right) \mathbf{v}_1 + c_2 \left(1 - \frac{\lambda_2}{\lambda_p}\right) \mathbf{v}_2 + \ldots + c_{p-1} \left(1 - \frac{\lambda_{p-1}}{\lambda_p}\right) \mathbf{v}_{p-1}$$

no $\frac{\lambda_j}{\lambda_p} = 1$ on RHS,

showed $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{p-1}\}$ is also dependent!!
Theorem 2

Let $A_{n \times n}$ have factored characteristic polynomial

$$p(\lambda) = (-1)^n (\lambda - \lambda_1)^{k_1} (\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_m)^{k_m}$$

where like terms have been collected so that each $\lambda_j$ is distinct (i.e., different). Notice that

$$k_1 + k_2 + \cdots + k_m = n$$

because the degree of $p(\lambda)$ is $n$.

- Then $1 \leq \dim(E_{\lambda = \lambda_j}) \leq k_j$. If $\dim(E_{\lambda = \lambda_j}) < k_j$ then the $\lambda_j$ eigenspace is called defective.
- The matrix $A$ is diagonalizable if and only if each $\dim(E_{\lambda = \lambda_j}) = k_j$. In this case, one obtains an $\mathbb{R}^n$ eigenbasis simply by combining bases for each eigenspace into one collection of $n$ vectors. (The same definitions and reasoning can apply to complex eigenvalues and eigenvectors, and a basis of $\mathbb{C}^n$.)

(The proof of this theorem is fairly involved. It was illustrated in a positive way by Exercise 2, and in a negative way by Exercise 3.)

\begin{equation*}
\text{yesterday} \quad B \quad \text{that was diagonalizable.} \quad |B - \lambda I| = -(\lambda - 2)^2 (\lambda - 3) \quad \dim E_{\lambda = 2} = 2 \quad \dim E_{\lambda = 1} = 1
\end{equation*}

\begin{equation*}
\text{warming today} \quad C \quad \text{was not diagonalizable.} \quad |C - \lambda I| = -(\lambda - 2)^3 (\lambda - 3) \quad \dim E_{\lambda = 2} = 1 \quad \text{← defective}
\end{equation*}
Tues Mar 13

• 5.3 Diagonalizable matrices and Similar matrices.

Announcements:

⊙ Office hours today canceled.
⊙ I’ll try this afternoon to put up a practice test.

‘til 12:57

Warm-up Exercise:

Find all eigenvalues, and eigenspace basis for

$$C = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

for triangular matrix, diag:

- Eigenvalues: $C - \lambda I = \begin{bmatrix} 2-2 & 1 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{bmatrix}$
  $\det = (2-\lambda)(3-\lambda)$

- $E_{\lambda=2} = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$
- $E_{\lambda=3} = \text{span} \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$

- $E_{\lambda=2} = \text{Nul} (C - 2I)$
- $E_{\lambda=3} = \text{Nul} (C - 3I)$

- Algebraic multiplicity of $\lambda=2$ is 2
- $p(\lambda) = -(\lambda-2)(\lambda-3)$
- But $E_{\lambda=2}$ is only 1-dim. (\(\neq\))

we only have 2 independent eigenvectors, so, no basis of $\mathbb{R}^3$ made out eigenvectors for $C$. 
Continuing with the example from yesterday ...

If, for the matrix $A_{n \times n}$, there is a basis for $\mathbb{R}^n$ consisting of eigenvectors of $A$, then we can understand the geometry of the transformation

$$T(\mathbf{x}) = A \mathbf{x}$$

almost as well as if $A$ is a diagonal matrix, and so we call such matrices \textit{diagonalizable}. Having such a basis of eigenvectors for a given matrix is also extremely useful for algebraic computations, and will give another reason for the word \textit{diagonalizable} to describe such matrices.

Use an $\mathbb{R}^3$ basis made of out eigenvectors of the matrix $B$ in Exercise 2, yesterday, and put them into the columns of a matrix we will call $P$. We could order the eigenvectors however we want, but we'll put the $E_{\lambda=2}$ basis vectors in the first two columns, and the $E_{\lambda=3}$ basis vector in the third column:

$$P := \begin{pmatrix}
0 & 1 & 1 \\
1 & 0 & 1 \\
2 & -2 & 1 \\
\end{pmatrix}.$$

Now do algebra (check these steps and discuss what's going on!)

$$B = \begin{pmatrix} 4 & -2 & 1 \\ 2 & 0 & 1 \\ 2 & -2 & 3 \end{pmatrix} \quad \text{and} \quad P := \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 2 & -2 & 1 \end{pmatrix}.$$

In other words,

$$B P = P D,$$

$$B = P D P^{-1}, \quad \text{mult by } P^{-1} \text{ on the right} \quad \text{mult by } P \text{ on the left}.$$
where D is the diagonal matrix of eigenvalues (for the corresponding columns of eigenvectors in P).

Equivalently (multiply on the right by $P^{-1}$ or on the left by $P^{-1}$):

$$B = PD P^{-1} \quad \text{and} \quad P^{-1}BP = D.$$

**Exercise 1**) Use one of the identities above to show how $B^{100}$ can be computed with only two matrix multiplications!

$$B^{100} = \underbrace{PD P^{-1} PDP^{-1} \cdots PDP^{-1}}_{100 \text{ times}} = PD^{100} P^{-1}$$

\[
\begin{bmatrix}
4 & -2 & 1 \\
2 & 0 & 1 \\
2 & -2 & 3 \\
\end{bmatrix}^{100}
\]

\[
\begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
2 & -2 & 1 \\
\end{bmatrix}^{100}
\]

\[
\begin{bmatrix}
\lambda_1 & 0 & 0 \\
0 & \lambda_2 & 0 \\
0 & 0 & \lambda_3 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
\lambda_1 & 0 & 0 \\
0 & \lambda_2 & 0 \\
0 & 0 & \lambda_3 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
\end{bmatrix}
\]
Definition: Let $A_{n \times n}$. If there is an $\mathbb{R}^n$ (or $\mathbb{C}^n$) basis $v_1, v_2, \ldots, v_n$ consisting of eigenvectors of $A$, then $A$ is called diagonalizable. This is precisely why:

Write $A v_j = \lambda_j v_j$ (some of these $\lambda_j$ may be the same, as in the previous example). Let $P$ be the matrix

$$P = [v_1 \mid v_2 \mid \ldots \mid v_n].$$

Then, using the various ways of understanding matrix multiplication, we see

$$A P = A [v_1 \mid v_2 \mid \ldots \mid v_n] = [\lambda_1 v_1 \mid \lambda_2 v_2 \mid \ldots \mid \lambda_n v_n]$$

$$= [v_1 \mid v_2 \mid \ldots \mid v_n] \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$ 

$$A P = P \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}.$$ 

Unfortunately, as we've already seen, not all matrices are diagonalizable:

Exercise 2) Show that

$$C := \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

is not diagonalizable. (Even though it has the same characteristic polynomial as $B$, which was diagonalizable.

This was your Tuesday warm-up exercise.
Similar matrices. This generalizes the way in which diagonalizable matrices are similar to diagonal ones:

**Definition** The \( n \times n \) matrices \( A, B \) are said to be *similar* if there is and invertible matrix \( P \) so that
\[
P^{-1} A P = B.
\]

Notice that *being similar* is an equivalence relation:

1) If \( A \) is similar to \( B \) with the matrix \( P \), then \( B \) is similar to \( A \), with the matrix \( P^{-1} \):
\[
P^{-1} A P = B \quad \Rightarrow \quad A = P B P^{-1}.
\]

2) \( A \) is similar to itself, with \( P = I \):
\[
A = I^{-1} A I
\]

3) Being similar is transitive: if \( A \) is similar to \( B \) and \( B \) is similar to \( C \), then \( A \) is similar to \( C \): If we have invertible matrices \( P, Q \) so that
\[
P^{-1} A P = B
\]
\[
Q^{-1} B Q = C
\]

then
\[
Q^{-1} P^{-1} A P Q = Q^{-1} B Q = C.
\]

so \( A \) is similar to \( C \) via the matrix \( PQ \).

These three "equivalence relations" mean that the space all \( n \times n \) matrices can be partitioned into subsets of matrices which are similar to each other.

We'll see tomorrow that similar matrices represent the same linear transformation from \( \mathbb{R}^n \) to \( \mathbb{R}^n \), but with the matrices expressed with respect to different bases. For now (and for one of your homework problems tomorrow), we need to know that

**Theorem** Let \( A \) and \( B \) be similar matrices. Then they have the same characteristic polynomial, so the same eigenvalues. (They won't have the same eigenvectors, though.)

**proof** Let
\[
P^{-1} A P = B.
\]

Then
\[
det(B - \lambda I) = det(P^{-1} A P - \lambda I)
\]
\[
= det(P^{-1} A P - \lambda P^{-1} I P)
\]
\[
= det(P^{-1} (A - \lambda I) P)
\]
\[ \det(P^{-1}) \det(A - \lambda I) \det(P) = \det(A - \lambda I). \]

QED