More on the diagonalization problem

Before Thanksgiving we discussed the diagonalization problem as a special case of the similarity problem for matrices.

A matrix $\mathbf{A}$ is diagonalizable if there exists a diagonal matrix $\mathbf{D}$ and an invertible matrix $\mathbf{P}$ such that

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

Two square matrices $\mathbf{A}$ and $\mathbf{B}$ are similar if there exists an invertible matrix $\mathbf{P}$ such that

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

Theorem. Suppose that $\mathbf{A}$ and $\mathbf{B}$ are similar matrices. Then $\mathbf{A}$ and $\mathbf{B}$

1. have the same characteristic polynomial and consequently the same eigenvalues, and
2. the same geometric multiplicities for each eigenvalue.

For an arbitrary matrix $\mathbf{A}$, what can be said about it if it is diagonalizable?

Example. Consider the matrix

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

Can it be diagonalized?

Example. Consider the matrix

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

Can it be diagonalized?

Let's see why having a basis of eigenvectors is enough to be able to diagonalize $\mathbf{A}$: Suppose $\mathbf{A}$ has n linearly independent eigenvectors

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

and let λ_i be the eigenvalue that is associated to $\mathbf{v}_i$. (Note: The λ_i need not be distinct.)

Then we can diagonalize $\mathbf{A}$ using the matrix

$$\mathbf{P} = \left[\begin{array}{c|c|c|c} \mathbf{v}_1 & \mathbf{v}_2 & \quad \cdots \quad & \mathbf{v}_n \end{array} \right].$$

Example. Let

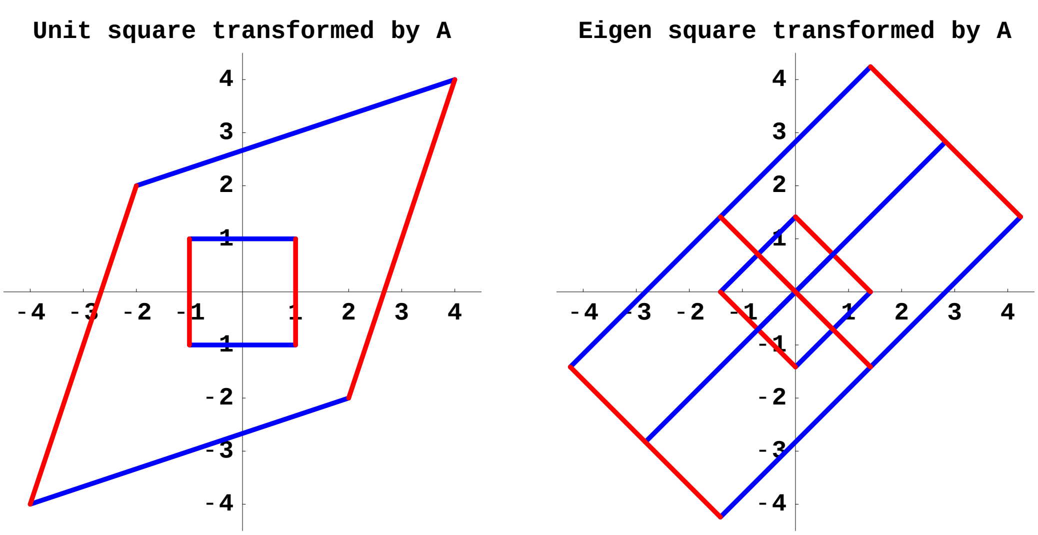
$$\mathbf{A} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

We have already seen that $\mathbf{A}$ has eigenvalues $\lambda = 2$ and $\lambda = 4$. The $\lambda = 4$ eigenspace is

$$\text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\},$$

and the $\lambda = 2$ eigenspace is

$$\text{Span}\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}.$$



Now let's return to the unusual matrix that is in the animation.

Example. Consider the matrix

$$\mathbf{A} = \begin{bmatrix} \frac{158}{165} & \frac{19}{495} \\ \frac{38}{165} & \frac{1043}{990} \end{bmatrix}$$

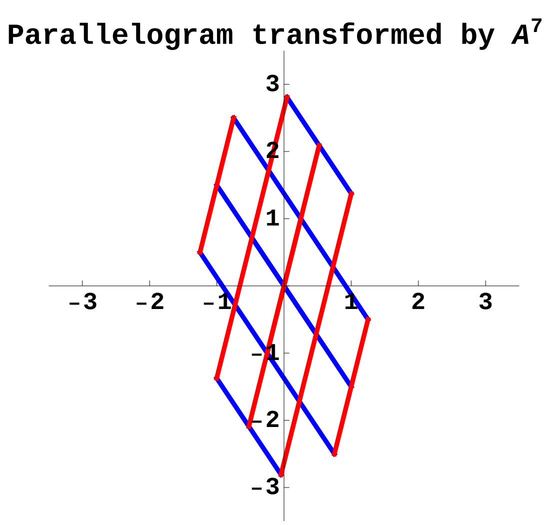
The vector

$$\begin{bmatrix} 1 \\ -\frac{3}{2} \end{bmatrix}$$

is an eigenvector corresponding to the eigenvalue $\lambda = 9/10$, and the vector

$$\begin{bmatrix} \frac{1}{4} \\ 1 \end{bmatrix}$$

is an eigenvector corresponding to the eigenvalue $\lambda = 10/9$.



Example. Let

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

We have already seen that $\mathbf{A}$ has eigenvalues $\lambda = 1$ and $\lambda = 2$. The $\lambda = 1$ eigenspace is the plane $x_1 + x_2 - x_3 = 0$, and the $\lambda = 2$ eigenspace is the line $x_1 = x_2 = x_3$.

The dot product

Definition. Given

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \qquad \text{and} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

in $\mathbb{R}^n$, then

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n.$$

Example. Let

$$\mathbf{u} = \begin{bmatrix} 1 \\ \sqrt{2} \\ \pi \end{bmatrix} \qquad \text{and} \qquad \mathbf{v} = \begin{bmatrix} 5 \\ -3\sqrt{2} \\ 2 \end{bmatrix}.$$

Remarks.

1. The dot product $\mathbf{u} \cdot \mathbf{v}$ can be viewed as matrix multiplication, that is, $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$.
2. Note that $\mathbf{u} \cdot \mathbf{v} \cdot \mathbf{w}$ is undefined.
3. Note that $(\mathbf{u} \cdot \mathbf{v})\mathbf{w}$ involves scalar multiplication.

Let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{B}$ be an $n \times p$ matrix. Write $\mathbf{A}$ and $\mathbf{B}$ as

$$\mathbf{A} = \begin{bmatrix} \mathbf{r}_1 \\ \hline \mathbf{r}_2 \\ \hline \vdots \\ \hline \mathbf{r}_m \end{bmatrix} \qquad \text{and} \qquad \mathbf{B} = \left[\begin{array}{c|c|c|c} \mathbf{c}_1 & \mathbf{c}_2 & \ldots & \mathbf{c}_p \end{array} \right].$$

Let's interpret the product $\mathbf{AB}$ in terms of the dot product.

Theorem 1. Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$, and let c be a scalar. Then

(a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$

(b) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$

(c) $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$

(d) $\mathbf{u} \cdot \mathbf{u} \geq 0$, and

(e) $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$

Definition. The length (or norm) of a vector $\mathbf{v}$ in $\mathbb{R}^n$ is

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

Note that $||r\mathbf{v}|| = |r|\,||\mathbf{v}||$.

Given $\mathbf{v} \neq \mathbf{0}$, we normalize $\mathbf{v}$ by computing the vector $\mathbf{u} = \dfrac{1}{||\mathbf{v}||}\mathbf{v}$.

If we think of two vectors $\mathbf{u}$ and $\mathbf{v}$ as points in $\mathbb{R}^n$, then we define the distance between $\mathbf{u}$ and $\mathbf{v}$ as

$$\text{dist}(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||.$$

What about angles? Let's start with right angles.

Definition. Two vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if $\mathbf{u}\cdot\mathbf{v} = 0$.