§5.2 Diagonalization

13 Apr 2017

We've learned that if an $n \times n$ matrix A has a collection of eigenvectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ that forms a basis for $\mathbb{R}^n$, then A acts simply in that basis:

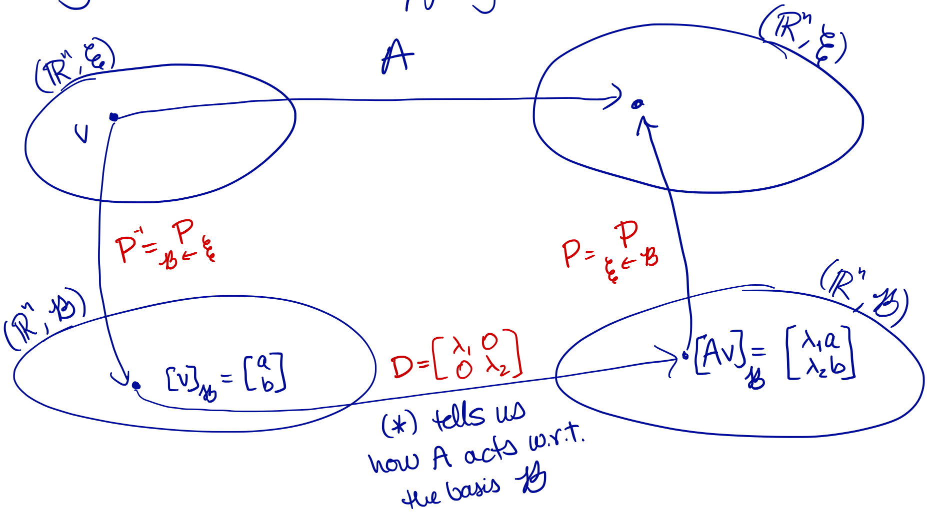
$$A(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) = c_1\lambda_1\vec{v}_1 + \cdots + c_n\lambda_n\vec{v}_n. \quad (*)$$

This is great for by-hand analysis, as we saw in our Fibonacci problem. But to make full computational use of this idea, we need to "matrixify" it. In other words, how does A act on a vector expressed in terms of its coordinates with respect to the standard basis?

Example Find a matrix formula for $A^n v$ based on $(*)$, where $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$.

Solution We have already seen that

$\lambda_1 = -4$, $\lambda_2 = 7$ are the eigenvalues of this matrix, with eigenvectors $\begin{bmatrix} -6 \\ 5 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectively. So we set $\mathcal{B} = \left\{ \begin{bmatrix} -6 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$, and apply A to v by first mapping to $\mathcal{B}$ coordinates, then using (*), then mapping back:



Recall that the columns of P are the basis vectors v_1, v_2. So we can see that

$$A = PDP^{-1}. \quad \text{Similarly, } A^n = PD^nP^{-1} = \boxed{P \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} P^{-1}}.$$

Example Show that $(PDP^{-1})^n = PD^nP^{-1}$ directly.

Solution

$$(PD\cancel{P^{-1})(P}D\cancel{P^{-1})\cdots(P}DP^{-1})$$

$$= PD^nP^{-1}.$$

Example Show directly from the defⁿ of eigenvectors/values that if A has n lin. ind. eigenvectors $\{v_1, \ldots, v_n\}$ with eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$, then $A = PDP^{-1}$ where $P = [\vec{v}_1 \cdots \vec{v}_n]$ and $D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ 0 & \cdots & & \lambda_n \end{bmatrix}$.

Solution $A = PDP^{-1} \iff AP = PD$, and

$$AP = A[\vec{v}_1 \cdots \vec{v}_n] \overset{\text{def}^n \text{ of matrix-matrix product}}{=} [A\vec{v}_1 \cdots A\vec{v}_n]$$

$$= [\lambda_1\vec{v}_1 \cdots \lambda_n\vec{v}_n]$$

again, def of matrix-matrix product

$$= [\vec{v}_1 \cdots \vec{v}_n]\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = PD. \quad \blacksquare$$

We say that a matrix A is **diagonalizable** if there exist P, D so that P is invertible, D is diagonal, and $A = PDP^{-1}$.

Diagonalizability is nice because, as we've seen, raising a matrix in diagonalized form to a power is very convenient. Actually, diagonalizability is intimately related to eigenvectors:

Theorem A is diagonalizable iff A possesses n linearly independent eigenvectors

Proof We've shown the $\Leftarrow$ direction. Now suppose A is diagonalizable, i.e., $A = PDP^{-1}$ for P invertible & D diagonal. Then $AP = PD$. If we write $P = [\vec{p}_1 \cdots \vec{p}_n]$ in terms of its columns and $D = \begin{bmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{bmatrix}$, then $AP = PD$ says that

$$A\vec{p}_j = d_j \vec{p}_j$$

for all $j = 1, \dots, n$. In other words, $\vec{p}_1, \dots, \vec{p}_n$ are eigenvectors of A. They're lin. ind. because they're column vectors of an invertible matrix. $\blacksquare$

Example Show that $A=\begin{bmatrix}1&1\\0&1\end{bmatrix}$ is not diagonalizable.

Solution The eigenvalues of A are the solutions of $\begin{vmatrix}1-\lambda & 1\\ 0 & 1-\lambda\end{vmatrix} = (1-\lambda)^2 = 0$, so just $\lambda=1$.

The eigenspace of A corresponding to $\lambda=1$ is the solution set of

$$\begin{bmatrix}0&1\\0&0\end{bmatrix}\vec{v}=\begin{bmatrix}0\\0\end{bmatrix}$$

which is $\left\{\begin{bmatrix}v_1\\0\end{bmatrix} : v_1\in\mathbb{R}\right\} = \mathrm{Span}\left(\left\{\begin{bmatrix}1\\0\end{bmatrix}\right\}\right)$. So there are no other eigenvectors of A other than $\begin{bmatrix}1\\0\end{bmatrix}$ & multiples thereof. So A cannot be diagonalizable.

Example Show that $\begin{bmatrix}1&0\\0&1\end{bmatrix}$ is diagonalizable.

Solution: it's already diagonal, silly!

** These examples show that repeated eigenvalues may or may not lead to diagonalyability failure. **

In other words, the eigenspaces of A might or might not "fill out" the whole space $\mathbb{R}^n$, meaning that their dimensions might sum to n (in which case A is diagonalizable), or something less (non-diagonalizable).

§5.4 Matrix similarity

Let T_1 be the $90°$ ccw rotation in $\mathbb{R}^2$, and T_2 the $90°$ cw rotation in $\mathbb{R}^2$. Clearly $T_1 \neq T_2$. However, they are geometrically similar, intuitively. More precisely,

$$T_1(x\vec{e}_1 + y\vec{e}_2) = -y\vec{e}_1 + x\vec{e}_2 \text{ , while}$$

$$T_2(x\vec{b}_1 + y\vec{b}_2) = -y\vec{b}_1 + x\vec{b}_2$$

where $\vec{b}_1 = \vec{e}_1$ and $\vec{b}_2 = -\vec{e}_2$. In other words,

(i.e. both work by applying $(x,y) \mapsto (-y,x)$ to coordinates with respect to some basis)

they're doing exactly the same thing coordinate-wise, just with respect to different bases.

In matrix terms, we can write T_2 as a comp. of 3 operations: change basis to $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$, apply T_1 to the new <u>coordinate</u> vector, then change basis back. I.e.:

standard basis

$$
\begin{tikzcd}[column sep=huge, row sep=huge]
(\mathbb{R}^2, \mathcal{E}) \arrow[r, bend left, "T_1(\vec{x}) = \underbrace{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}}_{A}"'] \arrow[d, "P^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}"'] & (\mathbb{R}^2, \mathcal{E}) \\
(\mathbb{R}^2, \mathcal{B}) \arrow[r, "T_2([\vec{x}]_{\mathcal{B}}) = \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}}_{B}[\vec{x}]_{\mathcal{B}}"'] & (\mathbb{R}^2, \mathcal{B}) \arrow[u, "P = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}"']
\end{tikzcd}
$$

alternatative basis vectors

$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$

If $A = PBP^{-1}$, we say A, B are similar.

So, a matrix is diagonalizable iff it's similar to a diagonal matrix.

☆ <u>Note</u>: row equivalence & similarity are <u>very</u> different. Eg. similar matrices have the same eigenvalues