18.06.26: More on similarity and diagonalizability

Lecturer: Barwick

Shadows are harshest
when there is only one lamp.
— James Richardson

Let's get back to *similarity*.

Suppose I have a basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of $\mathbf{R}^n$, and suppose A is an $n \times n$ matrix, giving us a linear map $T_A \colon \mathbf{R}^n \longrightarrow \mathbf{R}^n$.

Maybe T_A is actually more interesting to us than A, and maybe $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a better basis for us than the standard basis. So we want to express the action of T_A entirely in terms of $\{\vec{v}_1, \ldots, \vec{v}_n\}$.

When we look at our chosen basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$, we can write each vector $T_A(\vec{v}_j)$ in a unique fashion as a linear combination of the basis vectors:

$$T_A(\vec{v}_j) = \sum_{i=1}^{n} \beta_{ij}\vec{v}_i.$$

We could have put all those coefficients together into a new matrix

$$B = (\beta_{ij}).$$

We say that B *represents* T_A *with respect to the basis* $\{\vec{v}_1, \ldots, \vec{v}_n\}$.

If we'd done that with the standard basis, we'd have the matrix A staring back at us. But with a different basis, B isn't A. So how do they relate??

So, let's make a nice invertible matrix out of our basis:

$$V := \begin{pmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{pmatrix}.$$

We see that

$$AV = \begin{pmatrix} \sum_{i=1}^n \beta_{i1}\vec{v}_i & \cdots & \sum_{i=1}^n \beta_{in}\vec{v}_i \end{pmatrix}.$$

On the other hand,

$$VB = \begin{pmatrix} \sum_{i=1}^n \beta_{i1}\vec{v}_i & \cdots & \sum_{i=1}^n \beta_{in}\vec{v}_i \end{pmatrix}.$$

So $AV = VB$, whence $B = V^{-1}AV$.

18.06.26: More on similarity and diagonalizability

This is a tricky concept. I like to think about this diagram:

$$\begin{array}{ccc}
\mathbf{R}^n_{\hat{e}_i} & \xrightarrow{\quad A \quad} & \mathbf{R}^n_{\hat{e}_i} \\
{\scriptstyle V}\big\uparrow & & \big\downarrow{\scriptstyle V^{-1}} \\
\mathbf{R}^n_{\vec{v}_i} & \xrightarrow[\quad B \quad]{} & \mathbf{R}^n_{\vec{v}_i}
\end{array}$$

Definition. We say two $n \times n$ matrices A and B are *similar* if they represent the same linear transformation with respect to two different bases.

Equivalently, A and B are similar if B represents T_A with respect to some basis.

Equivalently, A and B are similar if and only if there is some invertible matrix V such that

$$B = V^{-1} A V.$$

Let's do a quick example. Consider $A = \begin{pmatrix} 5 & -3 \\ 2 & -2 \end{pmatrix}$, and let's write the matrix B that represents T_A with respect to the basis $\left\{ \vec{v}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}$.

$$
\begin{aligned}
B &= \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 5 & -3 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \\
&= \frac{1}{5} \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 5 & -3 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \\
&= \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}.
\end{aligned}
$$

So our A is actually similar to a diagonal matrix.

And what does that mean? The matrix B that represents A with respect to $\{\vec{v}_1, \vec{v}_2\}$ is $\operatorname{diag}(4, -1)$, so:

$$A\vec{v}_1 = 4\vec{v}_1 \quad \text{and} \quad A\vec{v}_2 = -\vec{v}_2.$$

So the eigenvalues for the diagonal matrix B are also eigenvalues for the matrix A.

In other words, $\vec{v}_1$ and $\vec{v}_2$ form a *basis of eigenvectors* for A. So A is *diagonalizable*.

In fact, similar matrices always have the same eigenvalues, because they have the same characteristic polynomials:

$$\begin{aligned} p_{V^{-1}AV}(t) = \det(tI - V^{-1}AV) &= \det(tV^{-1}V - V^{-1}AV) \\ &= \det(V^{-1}(tI - A)V) \\ &= (\det V)^{-1} \det(tI - A) \det V \\ &= \det(tI - A) = p_A(t). \end{aligned}$$

So now we understand our terminology: ***an*** $n \times n$ ***matrix*** A ***is diagonalizable if and only if it is similar to a diagonal matrix*** $\mathbf{diag}(\lambda_1, \ldots, \lambda_n)$***, where*** $\lambda_1, \ldots, \lambda_n$ ***are the eigenvalues of*** A ***with multiplicity.***

In other words, the following are logically equivalent for an $n \times n$ matrix A:

(1) A is diagonalizable.

(2) There exists a basis for $\mathbf{R}^n$ consisting of eigenvectors of A.

(3) There is a basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ for $\mathbf{R}^n$ such that the matrix that represents T_A with respect to $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is diagonal.

(4) A is similar to a diagonal matrix.

(5) A is similar to the diagonal matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, where the λ_i's are the eigenvalues of A, taken with multiplicity.

(6) There is an invertible $n \times n$ matrix V such that $\mathrm{diag}(\lambda_1, \ldots, \lambda_n) = V^{-1}AV$.

There's one more condition I'd like to add to this list. To describe it, we need some notation, which may work in unfamiliar way: suppose V, W, X are three vector subspaces of $\mathbf{R}^n$, and suppose $V \subseteq X$ and $W \subseteq X$. Then we write

$$X = V \oplus W$$

if every vector $\vec{x} \in X$ can be written *uniquely* as a sum $\vec{v} + \vec{w}$ with $\vec{v} \in V$ and $\vec{w} \in W$.

Equivalently, $X = V \oplus W$ if $V \cap W = \{0\}$ and if every vector $\vec{x} \in X$ can be written as a sum $\vec{v} + \vec{w}$.

In other words, if $V \cap W = \{0\}$, then

$$V \oplus W = \{\vec{x} \in \mathbf{R}^n \mid \vec{x} = \vec{v} + \vec{w}, \text{ where } \vec{v} \in V \text{ and } \vec{w} \in W\}.$$

The important fact here is that

$$\dim(V \oplus W) = \dim(V) + \dim(W).$$

That's because I can take a basis $\{\vec{v}_1, \ldots, \vec{v}_k\}$ of V and a basis $\{\vec{w}_1, \ldots, \vec{w}_\ell\}$ of W, and I can put them together into a basis

$$\{\vec{v}_1, \ldots, \vec{v}_k, \vec{w}_1, \ldots, \vec{w}_\ell\}.$$

So, in fact, if V, W, X are three vector subspaces of $\mathbf{R}^n$ with $V \subseteq X$ and $W \subseteq X$, then $X = V \oplus W$ if and only if: (1) $V \cap W = \{0\}$ and (2) $\dim V + \dim W = \dim X$.

Note that if λ and μ are two different eigenvalues of an $n \times n$ matrix A, then $L_\lambda \cap L_\mu = \{0\}$; indeed, if $\vec{v} \in L_\lambda \cap L_\mu$, then it is an eigenvector for both λ and μ. So

$$\lambda\vec{v} = A\vec{v} = \mu\vec{v}.$$

Thus $(\lambda - \mu)\vec{v} = \vec{0}$, and since $\lambda - \mu \neq 0$, we may divide by it to see that $\vec{v} = \vec{0}$.

Now we can add the last of our equivalent conditions for A to be diagonalizable:

(7) If $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of A, then

$$\mathbf{R}^n = L_{\lambda_1} \oplus \cdots \oplus L_{\lambda_k}.$$

(The cool kids call this the *spectral decomposition*.)

Proposition. *An $n \times n$ matrix with n distinct real eigenvalues is diagonalizable.*

Proof. Let $\lambda_1, \ldots, \lambda_n$ be the distinct eigenvalues. Let's look at the corresponding eigenspaces

$$L_{\lambda_i} = \ker(\lambda_i I - A),$$

each of which has $\dim(L_{\lambda_i}) \geq 1$.

We have already seen that if $i \neq j$, then $L_{\lambda_i} \cap L_{\lambda_j} = \{0\}$.

So we have $\mathbf{R}^n$, which is n-dimensional, and we have n different subspaces L_{λ_i}, each of which has dimension ≥ 1, and no two of which intersect nontrivially. So

$$\dim(L_{\lambda_1} \oplus \cdots \oplus L_{\lambda_n}) = \dim(L_{\lambda_1}) + \cdots + \dim(L_{\lambda_n}) \leq n.$$

But the only way for that to happen is if each $\dim(L_{\lambda_i}) = 1$, in which case their sum is exactly n. Hence

$$\mathbf{R}^n = L_{\lambda_1} \oplus \cdots \oplus L_{\lambda_n},$$

and so A is diagonalizable. □

So let's think again about our two obstructions to diagonalizability of A:

(1) Non-real eigenvalues.

(2) Repeated eigenvalues with an undersized eigenspace.

Spectral theorems are how we deal with point (2). We just proved one: *an* $n \times n$ *matrix with* n *distinct real eigenvalues is diagonalizable over* $\mathbf{R}$. Next time, we'll prove another: *a symmetric matrix is diagonalizable over* $\mathbf{R}$. Eventually, we'll pass to the complex numbers, and do linear algebra there.