18.06.30: Spectral theorem

Lecturer: Barwick

And in the telling of that story
I lose my way inside a prepositional phrase.
– Wye Oak

For $v, w \in \mathbf{C}^n$, write

$$\langle v|w\rangle := v^* w.$$

This is a complex number, called the *inner product* of two complex vectors; it extends the usual dot product, but notices that the linearity in the first coordinate is *twisted*:

$$\langle \alpha v|w\rangle = \overline{\alpha}\langle v|w\rangle \text{ but } \langle v|\alpha w\rangle = \alpha\langle v|w\rangle.$$

The length of a vector $v \in \mathbf{C}^n$ is defined by $||v||^2 = \langle v|v\rangle$; it's precisely the same as the length of the corresponding vector in $\mathbf{R}^{2n}$. (Why??)

Lemma. *An $n \times n$ complex matrix B is Hermitian if and only if, for any $v, w \in \mathbf{C}^n$,*

$$\langle Av|w\rangle = \langle v|Aw\rangle.$$

Proof. If A is Hermitian, then $(Av)^* w = v^* A^* w = v^* Aw$.

On the other hand, suppose that for any $v, w \in \mathbf{C}^n$,

$$\langle Av|w\rangle = \langle v|Aw\rangle.$$

Then when $v = \hat{e}_i$ and $w = \hat{e}_j$, this equation becomes

$$\overline{a}_{ji} = (A^i)^* \hat{e}_j = \hat{e}_i^* A^j = a_{ij}.$$

□

Theorem (Spectral theorem; last big result of the semester). *Suppose B a Hermitian matrix. Then*

(1) The eigenvalues of B are real.

(2) There is an orthogonal basis of eigenvectors for B; in particular, B is diagonalizable over $\mathbf{C}$ (and even over $\mathbf{R}$ if B has real entries).

Proof. Let's first see why the eigenvalues of B must be real. Suppose $v \in \mathbf{C}^n$ an eigenvector of B with eigenvalue λ, so that $Bv = \lambda v$. Then,

$$\begin{aligned} \lambda||v||^2 = \lambda\langle v|v\rangle &= \langle v|\lambda v\rangle \\ &= \langle v|Bv\rangle \\ &= \langle Bv|v\rangle \\ &= \langle \lambda v|v\rangle \\ &= \overline{\lambda}\langle v|v\rangle = \overline{\lambda}||v||^2. \end{aligned}$$

Since $v \neq 0$, one has $||v|| \neq 0$, whence $\lambda = \overline{\lambda}$.

Now let's see about that orthogonal basis of eigenvectors. Using the Fundamental Theorem of Algebra, write the characteristic polynomial

$$p_B(t) = (t - \lambda_1) \cdots (t - \lambda_n),$$

where $\lambda_1, \ldots, \lambda_n \in \mathbf{C}$ are the roots of p_B. We may *not* assume that the λ_i's are distinct!!

Let's choose an eigenvector v_1 with eigenvalue λ_1, and consider the hyperplane

$$W_1 \coloneqq \{w \in \mathbf{C}^n \mid \langle v_1 | w \rangle = 0\}.$$

Note that for any $w \in W_1$, one has

$$\langle v_1 | Bw \rangle = \langle Bv_1 | w \rangle = \langle \lambda_1 v_1 | Bw \rangle = \lambda_1 \langle v_1 | w \rangle = 0,$$

so $Bw \in W_1$ as well. Hence the linear map $T_B \colon \mathbf{C}^n \longrightarrow \mathbf{C}^n$ restricts to a map $T_1 \colon W_1 \longrightarrow W_1$.

Select, temporarily, a $\mathbf{C}$-basis $\{w_2, \ldots, w_n\}$ of W_1. Then $\{v_1, w_2, \ldots, w_n\}$ is a $\mathbf{C}$-basis of $\mathbf{C}^n$, and writing T_B relative to this basis gives us a matrix

$$C_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & B_1 \end{pmatrix},$$

where B_1 is the $(n-1) \times (n-1)$ matrix that represents T_1 relative to the basis $\{w_2, \ldots, w_n\}$, and

$$p_{B_1} = (t - \lambda_2) \cdots (t - \lambda_n).$$

Now we run that same argument again with the $(n-1)\times(n-1)$ matrix B_1 in place of the $n\times n$ matrix B to get:

- an eigenvector $v_2 \in W_1$ with eigenvalue λ_2,
- the subspace $W_2 \subset W_1$ of vectors orthogonal to v_2,
- and an $(n-2)\times(n-2)$ matrix B_2 that represents T_B restricted to W_2.

Now we find that B is similar to

$$C_2 = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & B_2 \end{pmatrix}.$$

We repeat this argument repeatedly on each new B_i, each time getting:

- an eigenvector $v_{i+1} \in W_i$ with eigenvalue λ_{i+1},
- the subspace $W_{i+1} \subset W_i$ of vectors orthogonal to v_{i+1},
- and an $(n-i-1) \times (n-i-1)$ matrix B_{i+1} that represents T_B restricted to W_{i+1}.

At each stage, we find that B is similar to

$$C_{i+1} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_i & 0 \\ 0 & 0 & \cdots & 0 & B_{i+1} \end{pmatrix}.$$

This process eventually stops, when $i = n$. Then we're left with:

- eigenvectors $v_1, \dots, v_n$ with eigenvalues $\lambda_1, \dots, \lambda_n$,
- a string of subspaces

$$\mathbf{C}^n = W_0 \supset W_1 \supset W_2 \supset \cdots \supset W_n = \{0\},$$

with $v_{i+1} \in W_i$, and

$$W_{i+1} = \{w \in W_i \mid \langle v_{i+1} | w \rangle\},$$

- and a diagonal matrix $C_n = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ to which B is similar. □