Lecture 24:

Spectral Theorem

Let $F = \mathbb{R}$ or $\mathbb{C}$, and let V be a vector space over F. Let $\langle , \rangle : V \times V \to F$ be the standard inner product (dot product/conjugate dot product).

Today's BIG theorem:

> **Spectral theorem.**
> Let $A \in M_n(\mathbb{R})$ be a symmetric matrix. Then A is diagonalizable over $\mathbb{R}$, and there exists an orthonormal basis for $\mathbb{R}^n$ (with respect to the standard inner product) consisting of eigenvectors for A.

Non-examples. Let

$$X = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} 2 & -9 \\ 0 & -7 \end{pmatrix}.$$

Both are square matrices over $\mathbb{R}$, but neither are symmetric.

Since X is already in Jordan canonical form, we know it is not diagonalizable. The eigenspaces of Y are

$$V_2 = \left\{ \begin{pmatrix} a \\ 0 \end{pmatrix} \,\middle|\, a \in \mathbb{R} \right\} \quad \text{and} \quad V_{-7} = \left\{ \begin{pmatrix} b \\ b \end{pmatrix} \,\middle|\, b \in \mathbb{R} \right\}.$$

So Y *is* diagonalizable, but the eigenspaces aren't orthogonal:

$$(a, 0) \cdot (b, b) = ab,$$

and $ab = 0$ only if $a = 0$ or $b = 0$

(either way, one of those two vectors must be 0).

Spectral theorem. Let $A \in M_n(\mathbb{R})$ be a symmetric matrix. Then A is diagonalizable over $\mathbb{R}$, and there exists an orthonormal basis for $\mathbb{R}^n$ (with respect to the standard inner product) consisting of eigenvectors for A.

Example. Let

$$A = \begin{pmatrix} -1 & -1 & -2 \\ -1 & -1 & 2 \\ -2 & 2 & 2 \end{pmatrix}.$$

The characteristic polynomial of A is $p_A(x) = (4-x)(-2-x)^2$, and the eigenspaces are

$$V_4 = \left\{ \begin{pmatrix} -\frac{1}{2}a \\ \frac{1}{2}a \\ a \end{pmatrix} \middle| a \in \mathbb{R} \right\} \quad \text{and} \quad V_{-2} = \left\{ \begin{pmatrix} b+2c \\ b \\ c \end{pmatrix} \middle| b, c \in \mathbb{R} \right\}$$

Note that

$$(-\tfrac{1}{2}a, \tfrac{1}{2}a, a) \cdot (b+2c, b, c) = -\tfrac{1}{2}a(b+2c) + \tfrac{1}{2}ab + ac = 0,$$

for all $a, b, c \in \mathbb{R}$. Hence $V_4 \perp V_{-2}$.

. . . In particular,

$$\mathcal{B}_4 = \left\{ \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} = \mathbf{v}_1 \right\} \quad \text{and} \quad \mathcal{B}_{-2} = \left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \mathbf{v}_2, \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \mathbf{v}_3 \right\}$$

are othonormal bases of V_4 and V_{-2}, respectively, and $\mathcal{B} = \mathcal{B}_4 \sqcup \mathcal{B}_{-2}$ is an orthonormal basis of V. So

$$A = \mathrm{Rep}_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id}) \mathrm{Rep}_{\mathcal{B}}^{\mathcal{B}}(A) \mathrm{Rep}_{\mathcal{E}}^{\mathcal{B}}$$

$$\begin{pmatrix} -1 & -1 & -2 \\ -1 & -1 & 2 \\ -2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} \begin{pmatrix} -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ 2/\sqrt{6} & 0 & 1/\sqrt{3} \end{pmatrix}^{-1}.$$

More: Since $\mathcal{B}$ is orthonormal, if $P = \mathrm{Rep}_{\mathcal{B}}^{\mathcal{E}}(\mathrm{id}) = \begin{pmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | \end{pmatrix}$, then

$$P^t P = \begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_1 \cdot \mathbf{v}_2 & \mathbf{v}_1 \cdot \mathbf{v}_3 \\ \mathbf{v}_2 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_2 & \mathbf{v}_2 \cdot \mathbf{v}_3 \\ \mathbf{v}_3 \cdot \mathbf{v}_1 & \mathbf{v}_3 \cdot \mathbf{v}_2 & \mathbf{v}_3 \cdot \mathbf{v}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad \text{So } P^{-1} = P^t.$$

A matrix $P \in M_{n\times n}(\mathbb{R})$ is **orthogonal** if its columns form an orthonormal set in $\mathbb{R}^n$.

Lemma. $P \in M_{n\times n}(\mathbb{R})$ is orthogonal if and only if $P^{-1} = P^t$.

Pf. The (i, j)-entry of P^tP is $\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{i,j}$.

Spectral theorem. If $A \in M_n(\mathbb{R})$ is symmetric, then A is diagonalizable over $\mathbb{R}$. Namely, there exists a real diagonal matrix D and an orthogonal matrix P such that $A = PDP^t$.

Claim 1. The characteristic polynomial of A splits over $\mathbb{R}$ (and, thus, the eigenvalues of A are all real).

Proof. By the fundamental theorem of algebra, the characteristic polynomial splits over $\mathbb{C}$:

$$p_A(x) = \prod_{k=1}^{n}(\lambda_k - x)$$

with $\lambda_k \in \mathbb{C}$. We must show each $\lambda_k \in \mathbb{R}$.
Fix $\lambda = \lambda_k$ for some k and take nonzero $\mathbf{v} \in \mathbb{C}^n$ such that $A\mathbf{v} = \lambda\mathbf{v}$.

Aside: For $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$, the standard inner product is

$$\mathbf{a} \cdot \overline{\mathbf{b}} = \mathbf{a}^t\overline{\mathbf{b}} = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} \overline{b_1} \\ \vdots \\ \overline{b_n} \end{pmatrix}.$$

So for $X \in M_n(\mathbb{C})$,

$$\langle X\mathbf{a}, \mathbf{b}\rangle = (X\mathbf{a})^t\overline{\mathbf{b}} = \mathbf{a}^tX^t\overline{\mathbf{b}} = \mathbf{a}^t\overline{\left(\overline{X}^t\mathbf{b}\right)} = \langle \mathbf{a}, \overline{X}^t\mathbf{b}\rangle.$$

Now, specifically A is both real and symmetric. So $\overline{A}^t = A$. So for any $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, we have $\langle A\mathbf{a}, \mathbf{b}\rangle = \langle \mathbf{a}, A\mathbf{b}\rangle$. In particular, for the eigenvector $\mathbf{v}$ above, we have

$$\lambda\langle \mathbf{v}, \mathbf{v}\rangle = \langle \lambda\mathbf{v}, \mathbf{v}\rangle = \langle A\mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{v}, A\mathbf{v}\rangle = \langle \mathbf{v}, \lambda\mathbf{v}\rangle = \overline{\lambda}\langle \mathbf{v}, \mathbf{v}\rangle.$$

So since $\langle \mathbf{v}, \mathbf{v}\rangle \neq 0$, we have $\lambda = \overline{\lambda}$, and therefore $\lambda \in \mathbb{R}$. $\square$

Claim 2. There exists P orthogonal such that $A = PDP^t$ (where $D = \operatorname{diag}(\lambda_1, \dots, \lambda_k)$ from Claim 1).

Pf. If $n = 1$, any 1×1 matrix is diagonal already.

Now induct on n: let $n > 1$, and take an eigenvalue-eigenvector pair $\lambda_1 \in \mathbb{R}$ and $\mathbf{v}_1 \in \mathbb{R}^n$; without loss of generality, take $\mathbf{v}_1$ to be a unit vector (V_{λ_1} is a vector space, so we can scale). Since $\{\mathbf{v}_1\}$ is orthonormal, it extends to an (ordered) orthonormal basis for $\mathbb{R}^n$, $\langle\!\langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle\!\rangle$. Let Q be the orthogonal matrix with columns $\mathbf{v}_1, \dots, \mathbf{v}_n$, and define

$$\widetilde{A} = Q^{-1}AQ = Q^tAQ.$$

Consider the structure of $\widetilde{A}$, and use induction:

1. $\widetilde{A} = Q^tAQ$ is symmetric. (Compute $(Q^tAQ)^t$ and see what happens.)
2. $\widetilde{A}$ has the form

$$\widetilde{A} = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right) \text{ for some } B \in M_{n-1}(\mathbb{R}).$$

   Since $\widetilde{A}$ is real and symmetric, so is B. Induct!

By induction,

$$\widetilde{A} = \underbrace{\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & T & \\ 0 & & & \end{array}\right)}_{S} \underbrace{\left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & E & \\ 0 & & & \end{array}\right)}_{D} \underbrace{\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & T^t & \\ 0 & & & \end{array}\right)}_{S^t}$$

for some orthonormal T and real diagonal $E \in M_{n-1}(\mathbb{R})$.

We have $\widetilde{A} = Q^tAQ = SDS^t$ with Q and S orthogonal and D a real diagonal matrix.

Define $P = QS$. Then

1. P is orthogonal (compute P^tP and see what happens), and
2. $A = PDP^t$:

$$PDP^t = (QS)D(QS)^t = QSDS^tQ^t = Q\widetilde{A}Q^t = A.$$

You try: The matrix

$$A = \begin{pmatrix} 10 & 2 & -2 & 0 & 0 \\ 2 & 7 & -1 & 0 & 0 \\ -2 & -1 & 7 & 0 & 0 \\ 0 & 0 & 0 & 9 & -3 \\ 0 & 0 & 0 & -3 & 9 \end{pmatrix} \in M_5(\mathbb{R})$$

has eigenvalues 6 and 12; the corresponding eigenspaces in $V = \mathbb{R}^5$ are

$$V_6 = \{(a_1 - a_2, 2a_2, 2a_1, a_3, a_3)^t \mid a_i \in \mathbb{R}\} \quad \text{and}$$
$$V_{12} = \{(2b_1, b_1, -b_1, b_2, -b_2)^t \mid b_i \in \mathbb{R}\}$$

1. Verify that for all $\mathbf{u} \in V_6$ and $\mathbf{v} \in V_{12}$, we have $\mathbf{u} \cdot \mathbf{v} = 0$. Why do we care?
2. Pick a basis S of V_6. Perform Gram-Schmidt on S to get an orthogonal basis of V_6; then normalize to get an orthonormal basis $\mathcal{B}_6$ of V_6.
3. Pick a basis S' of V_{12}. Perform Gram-Schmidt on S' to get an orthogonal basis of V_{12}; then normalize to get an orthonormal basis $\mathcal{B}_{12}$ of V_{12}.
4. Verify that $\mathcal{B} = \mathcal{B}_6 \sqcup \mathcal{B}_{12}$ is an orthonormal basis of V.
   (How can you do this *without* row-reducing a matrix?)
5. Give an orthogonal P and diagonal D such that $A = PDP^{-1}$.

Spectral theorem for complex matrices

A matrix $A \in M_n(\mathbb{C})$ is **Hermitian** if $\overline{A}^t = A$.

A matrix $U \in M_{n\times n}(\mathbb{C})$ is **unitary** if its columns are orthonormal, or equivalently, if U is invertible with $U^{-1} = \overline{U}^t$.

Theorem. (Spectral theorem) Let $A \in M_n(\mathbb{C})$ be a Hermitian matrix. Then $A = UD\overline{U}^t$ where U is unitary and D is a real diagonal matrix.

Proof follow similarly as in the real case.