Sketch of Lecture 15

Mon, 2/12/2024

Review: Diagonalizability

Example 86. (review) If A is a 2×2 matrix with $\det(A) = -8$ and eigenvalue 4. What is the second eigenvalue?

Solution. Recall that $\det(A)$ is the product of the eigenvalues (see below). Hence, the second eigenvalue is -2.

$\det(A)$ is the product of the eigenvalues of A.

Why? Recall how we determine the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of an $n\times n$ matrix A. We compute the characteristic polynomial $\det(A-\lambda I)$ and determine the λ_i as the roots of that polynomial.

That means that we have the factorization $\det(A-\lambda I) = (\lambda_1-\lambda)(\lambda_n-\lambda)\cdots(\lambda_n-\lambda)$. Now, set $\lambda = 0$ to conclude that $\det(A) = \lambda_1\lambda_2\cdots\lambda_n$.

Lemma 87. A matrix A is diagonalizable if and only if, for every eigenvalue λ that is k times repeated, the λ-eigenspace of A has dimension k.

In short, an $n\times n$ matrix A is diagonalizable if and only if there exists a basis of $\mathbb{R}^n$ consisting of eigenvectors of A (i.e. "there are enough eigenvectors").

The next two examples illustrate that not all matrices are diagonalizable and that, even if a real matrix is diagonalizable, the eigenvalues and eigenvectors might be complex.

Example 88. What are the eigenvalues and eigenvectors of $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$? Is A diagonalizable?

Solution. The characteristic polynomial is $\det\left(\begin{bmatrix} -\lambda & 1 \\ 0 & -\lambda \end{bmatrix}\right) = \lambda^2$, which has $\lambda = 0$ as a double root.

However, the 0-eigenspace $\operatorname{null}(A) = \operatorname{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$ is only 1-dimensional.

As a consequence, A is not diagonalizable.

Example 89. What are the eigenvalues and eigenvectors of $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$? Is A diagonalizable?

Solution. The characteristic polynomial is $\det\left(\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix}\right) = \lambda^2 + 1 = (\lambda - i)(\lambda + i)$.

Hence, the eigenvalues are $\pm i$.

The i-eigenspace $\operatorname{null}\left(\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix}\right)$ has basis $\begin{bmatrix} i \\ 1 \end{bmatrix}$.

The $-i$-eigenspace $\operatorname{null}\left(\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix}\right)$ has basis $\begin{bmatrix} -i \\ 1 \end{bmatrix}$.

Thus, A has the diagonalization $A = PDP^{-1}$ with $D = \begin{bmatrix} i & \\ & -i \end{bmatrix}$ and $P = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$.

Armin Straub
straub@southalabama.edu

The spectral theorem

Recall that a matrix A is symmetric if and only if $A^T = A$.

Theorem 90. (spectral theorem, long version) Suppose A is a symmetric matrix.

- A is always diagonalizable.
- All eigenvalues of A are real.
- The eigenspaces of A are orthogonal.

Proof. We will prove (parts of) the spectral theorem later on. For now, we just appreciate that the spectral theorem guarantees all these nice things to happen for symmetric matrices (for any specific A we know how to determine whether A is diagonalizable and what its eigenspaces are).

Comment. The eigenspaces of A being orthogonal means that eigenvectors for different eigenvalues are always orthogonal.

Important consequence. In the diagonalization $A = PDP^{-1}$, we can choose P to be orthogonal (in which case $P^{-1} = P^T$). In that case, the diagonalization takes the special form $A = PDP^T$, where P is orthogonal and D is diagonal.

(spectral theorem, compact version) A symmetric matrix A can always be diagonalized as $A = PDP^T$, where P is orthogonal and D is diagonal (and both are real).

How? We proceed as in the diagonalization $A = PDP^{-1}$. For a symmetric matrix A, we can arrange P to be orthogonal, by normalizing its columns. If there is a repeated eigenvalue, then we also need to make sure to pick an orthonormal basis for the corresponding eigenspace (for instance, using Gram–Schmidt).

Advanced comment. A matrix such that $A^T A = AA^T$ is called **normal**. For normal matrices, the (complex!) eigenspaces are again orthogonal to each other. However, normal matrices which are not symmetric will always have complex eigenvalues. (In that case, the orthogonal matrix P gets replaced with a unitary matrix, the complex version of orthogonal matrices, and the P^T becomes the conjugate transpose $P^* = \bar{P}^T$.)

Example 91.

(a) Determine the eigenspaces of the symmetric matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$.

(b) Diagonalize A as $A = PDP^T$.

Solution.

(a) The characteristic polynomial is $\begin{vmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{vmatrix} = (\lambda - 4)(\lambda + 2)$, and so A has eigenvalues $4, -2$.

The 4-eigenspace is $\operatorname{null}\left(\begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix}\right)$ has basis $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

The -2-eigenspace is $\operatorname{null}\left(\begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}\right)$ has basis $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Important observation. The 4-eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and the -2-eigenvector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ are orthogonal!

Review. The product of all eigenvalues $-2 \cdot 4 = -8$ equals the determinant $\det(A) = 1 - 9 = -8$.

(b) Note that a usual diagonalization is of the form $A = PDP^{-1}$.

We need to choose P so that $P^{-1} = P^T$, which means that P must be **orthogonal** (meaning orthonormal columns). [Choosing such a P is only possible if the eigenspaces of A are orthogonal.]

Hence, we normalize the two eigenvectors to $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

With $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix}$, we then have $A = PDP^T$.

Example 92. (again, simplified) Diagonalize the symmetric matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$ as $A = PDP^T$.

Solution. See Example 91 for a solution that illustrates how to diagonalize any symmetric matrix. For a simplified solution, note that we can see right away that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a 4-eigenvector (since the row sums are equal!).

Because the eigenspaces are orthogonal (since A is symmetric!), $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ must also be an eigenvector.

Indeed, $\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ shows that the corresponding eigenvalues is -2.

We normalize the two eigenvectors and use them as the columns of P, so that $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ is an orthogonal matrix ($P^{-1} = P^T$). With $D = \begin{bmatrix} 4 & 0 \\ 0 & -2 \end{bmatrix}$ we then have $A = PDP^T$.

Example 93. Let A be a symmetric 2×2 matrix with 7-eigenvector $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$ and $\det(A) = -21$. Determine the second eigenvalue and a corresponding eigenvector.

Further, diagonalize A as $A = PDP^T$.

Solution. A has $-\frac{21}{7} = -3$-eigenvector $\begin{bmatrix} -5 \\ 2 \end{bmatrix}$.

Hence, $A = PDP^T$ with $D = \begin{bmatrix} 7 & \\ & -3 \end{bmatrix}$ and $P = \frac{1}{\sqrt{29}}\begin{bmatrix} 2 & -5 \\ 5 & 2 \end{bmatrix}$.

Comment. Recall that, because A is symmetric, the eigenvector must be orthogonal to $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$.

[In general, $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} -b \\ a \end{bmatrix}$ are orthogonal.]

Let us prove the following important part of the spectral theorem.

Theorem 94. If A is symmetric, then the eigenspaces of A are orthogonal.

Proof. To prove the claim we need to show that, if $\boldsymbol{v}$ and $\boldsymbol{w}$ are eigenvectors of A with different eigenvalues (say λ and μ), then $\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{0}$. Suppose therefore that $A\boldsymbol{v} = \lambda\boldsymbol{v}$ and $A\boldsymbol{w} = \mu\boldsymbol{w}$ with $\lambda \neq \mu$.

First, we observe that, for any matrix A and vectors $\boldsymbol{v}$, $\boldsymbol{w}$, we have the following:

$$(A\boldsymbol{v}) \cdot \boldsymbol{w} = (A\boldsymbol{v})^T\boldsymbol{w} = (\boldsymbol{v}^T A^T)\boldsymbol{w} = \boldsymbol{v}^T(A^T\boldsymbol{w}) = \boldsymbol{v} \cdot (A^T\boldsymbol{w})$$

If A is symmetric, we therefore have

$$(A\boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{v} \cdot (A\boldsymbol{w}).$$

We now use that $A\boldsymbol{v} = \lambda\boldsymbol{v}$ and $A\boldsymbol{w} = \mu\boldsymbol{w}$ to conclude from the above that $\lambda\boldsymbol{v} \cdot \boldsymbol{w} = \mu\boldsymbol{v} \cdot \boldsymbol{w}$.

However, since $\lambda \neq \mu$, this is only possible if $\boldsymbol{v} \cdot \boldsymbol{w} = 0$. □