Math 2270-1

Notes of 10/25/19

- Announcement: slightly changed schedule, check online

5.4-5.5 More on Eigenvalues and Eigenvectors

- All matrices are square. Unless stated otherwise the number of rows and columns equals n.
- Recall: A matrix A is **diagonalizable** if there exists a non-singular matrix P and a diagonal matrix D such that

$$D = P^{-1}AP.$$

- In that formula, P is the matrix of eigenvectors of A, and the diagonal entries of D are the corresponding eigenvalues.
- A matrix is diagonalizable if and only if it has a basis of eigenvectors.
- A matrix that is not diagonalizable is **defective**.

diagonalizability is separate from invertibility.

	singular	**invertible**
defective:	$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$
diagonalizable:	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Matrix Construction

- It is sometimes useful to be able to construct a matrix with given eigenvalues and eigenvectors. Note that

  $$D = P^{-1}AP$$

  is equivalent to

  $$A = PDP^{-1}.$$

  Suppose you want to construct a matrix A with given eigenvalues and given eigenvectors. Proceed as follows:

1. Collect the eigenvectors into the matrix P as before.
2. Compute P^{-1}.
3. Compute
$$A = PDP^{-1}. \tag{1}$$

The Jordan Canonical Form

- It is not always possible to diagonalize a matrix. However, for all matrices A there exists a similarity transform to its **Jordan Canonical Form**[−1−] (named after Camille Jordan, 1838-1922).
- The JCF is a block diagonal matrix

$$P^{-1}AP = \begin{bmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_k \end{bmatrix}$$

  where each diagonal block is of the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_i & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda_i & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda_i & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{bmatrix}$$

- Apart from reordering the diagonal blocks the JCF is unique.
- Each Jordan block J_i corresponds to one eigenvector with eigenvalue λ_i.
- A matrix is diagonalizable if and only if all of its Jordan blocks are 1×1.

[−1−] The textbook mentions the Jordan Canonical Form in a footnote on page 294.

- The **algebraic multiplicity** of an eigenvalue is its order as a root of the characteristic polynomial. Its **geometric multiplicity** is the dimension of its eigenspace.

- Here is an example. Suppose the Jordan form of a matrix is given by

$$J = \begin{bmatrix} 2 & 1 & . & . & . & . & . & . & . & . \\ . & 2 & . & . & . & . & . & . & . & . \\ . & . & 2 & . & . & . & . & . & . & . \\ . & . & . & 3 & 1 & . & . & . & . & . \\ . & . & . & . & 3 & 1 & . & . & . & . \\ . & . & . & . & . & 3 & . & . & . & . \\ . & . & . & . & . & . & 3 & . & . & . \\ . & . & . & . & . & . & . & 4 & . & . \\ . & . & . & . & . & . & . & . & 4 & . \\ . & . & . & . & . & . & . & . & . & 5 \end{bmatrix}$$

- Entries indicated by dots are zero.

- The characteristic polynomial of this matrix is

  $$p(\lambda) = |J - \lambda I| = (2-\lambda)^3(3-\lambda)^4(4-\lambda)^2(5-\lambda).$$

  The number 2 is an eigenvalue of algebraic multiplicity 3 and geometric multiplicity 2, 3 is an eigenvalue of algebraic multiplicity 3 and geometric multiplicity 2, 4 is an eigenvalue of algebraic and geometric multiplicity 2, and 5 is an eigenvalue of algebraic and geometric multiplicity 1. The dimension of the space spanned by all eigenvectors is the sum of the geometric multiplicities which is 7. The matrix is defective.

Distinct Eigenvalues

- A set of eigenvectors corresponding to distinct eigenvalues is linearly independent. (The word "distinct" means that no two of the eigenvalues are equal.)

- Suppose we have a set of k eigenvectors $\mathbf{x}_i$ with corresponding eigenvalues λ_i, i.e.,

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \qquad i = 1, 2, \ldots, k$$

and

$$i \neq j \qquad \Longrightarrow \lambda_i \neq \lambda_j.$$

- Suppose the eigenvectors are in fact linearly dependent. Then one of them, say $\mathbf{x}_1$, can be written as a linear combination of some of the others. By omitting unnecessary vectors, and relabeling vectors if necessary, we can find a set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ such that $\mathbf{x}_1$ can be written **uniquely** as

$$\mathbf{x}_1 = \sum_{j=2}^{m} \alpha_j \mathbf{x}_j.$$

This implies that the set $\{\mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is linearly independent. Then we get

$$\mathbf{0} = \mathbf{x}_1 - \sum_{j=2}^{m} \alpha_j \mathbf{x}_j.$$

Multiplying with A gives

$$\mathbf{0} = A\mathbf{0} = \lambda_1 \mathbf{x}_1 - \sum_{j=2}^{m} \alpha_j \lambda_j \mathbf{x}_j.$$

- We now consider two cases. If $\lambda_1 \neq 0$ we can write
$$\mathbf{x}_1 = \sum_{j=2}^{m} \alpha_j \frac{\lambda_j}{\lambda_1} \mathbf{x}_j.$$
Since the eigenvalues are distinct we get a different linear combination for $\mathbf{x}_1$ which contradicts our assumption that the coefficients of the linear combination are unique.

- If $\lambda_1 = 0$ we get the equation
$$\mathbf{0} = \sum_{j=2}^{m} \alpha_j \lambda_1 \mathbf{x}_j$$
which contradicts our assumption that the set $\{\mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is linearly independent.

- So the original set
$$\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$$
is linearly independent.

- Recall that a matrix is diagonalizable if it has a set of n linearly independent eigenvectors.
- Thus a matrix with distinct eigenvalues is diagonalizable.
- This implies, for example, that the JCF can be computed only in exact arithmetic.
- A non-diagonalizable matrix must have multiple eigenvalues.
- Using inexact arithmetic, for example floating point arithmetic, would introduce errors that are effectively random. As a result the eigenvalues would be generically distinct, and the JCF would be diagonal.
- If a matrix is not diagonalizable then it is possible to change its entries by an arbitrarily small amount and make the matrix diagonalizable.
- Conceptually this is similar to invertibility: generically zero is not an eigenvalue and the matrix is invertible.

Complex Eigenvalues

- We saw that the characteristic polynomial of a real matrix is a polynomial of degree n with real coefficients.
- A polynomial with real coefficients may have complex roots. If it does then those complex roots occur in conjugate complex pairs.
- We also saw that for every (suitably normalized) polynomial with real coefficients there exists a real matrix with that polynomial as its characteristic polynomial.
- So every polynomial (with leading term $(-\lambda)^n$) is the characteristic polynomial of infinitely many (all similarity transforms of the companion matrix) matrices.
- Moreover, complex eigenvalues can occur naturally in applications.
- Here is a very simple example. Suppose

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

  Thus the map $\mathbf{x} \longrightarrow A\mathbf{x}$ rotates $\mathbf{x}$ counterclockwise by an angle θ.
- Its eigenvalues can't be real (unless θ is an integer multiple of π).
- So the eigenvalues of A must be complex. Let's compute them.

The most important thing to know about complex eigenvalues is that **symmetric real matrices don't have any!** The textbook addresses this issue in problem 24 on page 303 (and later in chapter 7).

- But the argument is quite simple.
- For any matrix A or vector $\mathbf{x}$ let

$$A^H = \bar{A}^T \quad \text{and} \quad \mathbf{x}^H = \bar{\mathbf{x}}^T$$

  where the bar denotes conjugate complex.
- A complex matrix A is **Hermitian**[−2−] if

$$A = \bar{A}^T.$$

- We will show that the eigenvalues of a Hermitian matrix are real.

Note that symmetric real matrices are special cases of Hermitian matrices.

−2− named after Charles Hermite, 1822–1901.

- Suppose
$$A\mathbf{x} = \lambda \mathbf{x} \tag{2}$$
where $A = A^H$, and A, λ, and $\mathbf{x}$ are all possibly complex. Taking the conjugate complex on both sides turns this into
$$\mathbf{x}^H A^H = x^H A = \bar{\lambda}\mathbf{x}^H. \tag{3}$$
Left multiplying with $\mathbf{x}^H$ in (2) and right multiplying with $\mathbf{x}$ in (3) gives
$$\mathbf{x}^H A\mathbf{x} = \lambda \mathbf{x}^H \mathbf{x} \quad \text{and} \quad \mathbf{x}^H A\mathbf{x} = \bar{\lambda}\mathbf{x}^H \mathbf{x}.$$
Thus
$$\lambda \mathbf{x}^H \mathbf{x} = \bar{\lambda}\mathbf{x}^H \mathbf{x}.$$
This implies that $\lambda = \bar{\lambda}$, i.e., λ is real.
- This is a great example of simplifying a problem by generalizing it!