

Math 2270-1

Notes of 10/23/2019

- Throughout: all matrices are square. Unless otherwise specified, n is the number of rows and columns.

Review of Change of Basis

- Suppose the columns \mathbf{v}_i of the matrix

$$P = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_n]$$

form a Basis of \mathbb{R}^n . Then any vector \mathbf{x} in \mathbb{R}^n can be written uniquely as a linear combination of the columns of P :

$$\mathbf{x} = \sum_{i=1}^n \beta_i \mathbf{v}_i = P[\mathbf{x}]_V$$

where

$$[\mathbf{x}]_V$$

is the **coordinate vector of \mathbf{x} with respect to the basis**

$$V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$$

- Note that \mathbf{x} is actually the coordinate vector of \mathbf{x} with respect to the standard basis

$$E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}.$$

- To convert a coordinate vector to the standard form we simply multiply with P .
- To convert from the standard basis to the basis V we multiply with P^{-1} :

$$\mathbf{x} = [\mathbf{x}]_{\mathbf{E}} = P[\mathbf{x}]_V \quad \text{and} \quad [\mathbf{x}]_V = P^{-1}\mathbf{x} = P^{-1}[\mathbf{x}]_{\mathbf{E}}.$$

- Suppose now that we have the linear transformation

$$\mathbf{x} \longrightarrow A\mathbf{x}.$$

We want to express this in terms of the coordinate vectors with respect to the basis V :

$$[A\mathbf{x}]_V = T[\mathbf{x}]_V.$$

What is T ?

- We get

$$[A\mathbf{x}]_V = P^{-1}A\mathbf{x} = P^{-1}AP[\mathbf{x}]_V.$$

- Hence

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



T expresses the same linear transformation as A , just in terms of a different basis.

Review of Similarity

- **Definition:** Two matrices A and B are **similar** if there is a non-singular matrix P such that

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

- Similar matrices have the same eigenvalues, and their eigenvectors are related in a straightforward way. To see this suppose that

$$A\mathbf{x} = \lambda\mathbf{x}$$

and note that

$$B(P^{-1}\mathbf{x}) = P^{-1}APP^{-1}\mathbf{x} = P^{-1}\lambda\mathbf{x} = \lambda(P^{-1}\mathbf{x}).$$

- Thus $P^{-1}\mathbf{x}$ is an eigenvector of B with eigenvalue λ .
- Another way to see that similar matrices have the same eigenvalues is to observe that their characteristic polynomials are the same. Using the multiplicative property of determinants and the fact that the determinant of the inverse is the reciprocal of the determinant of the original matrix we see

$$\begin{aligned} |B - \lambda I| &= |P^{-1}AP - \lambda P^{-1}IP| \\ &= |P^{-1}(A - \lambda I)P| \\ &= |P^{-1}| |A - \lambda I| |P| \\ &= |A - \lambda I| \end{aligned}$$

$$(D\mathbf{x})_i = \lambda_i x_i$$

- We'll also use the following observation. If we multiply a vector \mathbf{x} with a diagonal matrix D then each entry of $D\mathbf{x}$ will depend only on the corresponding entry of \mathbf{x} . It will be independent of the other entries of \mathbf{x} . If \mathbf{x} is a coordinate vector with respect to some basis then this means the the coordinate with respect to any basis vector is processed independently of the coordinates with respect to the other basis vector.

5.3 Diagonalization

- **Definition:** A matrix is **diagonalizable** if it is similar to a diagonal matrix.
- In other words, A is diagonalizable if there exists a diagonal matrix D and a non-singular matrix P such that

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

- By the end of today's meeting we will understand the following facts:
1. A is diagonalizable if and only if its eigenvectors can be used to form a basis of \mathbb{R}^n . (Unsurprisingly, such a basis is called an **eigenvector basis**.)
 2. If A is diagonalizable the diagonal entries of D are the eigenvalues of A and the columns of P are the corresponding eigenvectors.
 3. Some matrices are not diagonalizable! They don't have a set of n linearly independent eigenvectors. We'll see examples below.
 4. Diagonalizing a matrix amounts to using its eigenvectors as a basis of \mathbb{R}^n .

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



Things get more complicated if eigenvalues are **complex**. The statements above remain essentially true, but we need to consider complex matrices. Rather than doing that, for the time being assume all eigenvalues and eigenvectors are real.

Point Number 4. above is actually the easiest to see if you remember our discussion of **change of basis**.

- Suppose we want to express the linear transformation $y = A\mathbf{x}$ in terms of some (not necessarily eigenvector) basis

$$V = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$$

with corresponding matrix

$$P = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_n]$$

- We saw that

$$[A\mathbf{x}]_V = P^{-1}AP[\mathbf{x}]_V.$$

where $[\mathbf{x}]_V$ is the coordinate vector of \mathbf{x} with respect to the basis V .

- If V is a basis of eigenvectors we get that $P^{-1}AP$ is the diagonal matrix of eigenvalues, and the entries of the coordinate vector of some vector \mathbf{x} do not interact when being multiplied with $D = P^{-1}AP$.

- Suppose

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{n-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{bmatrix}$$

and $\{v_1, \dots, v_n\}$

$$P = [v_1 \quad v_2 \quad \dots \quad v_n]$$

where the v_i are the columns of P .

- Rewrite

$$D = P^{-1}AP \quad (1)$$

as

$$AP = PD \quad (2)$$

$AP \neq DP$

- The left side of this equation equals

$$[Av_1 \quad Av_2 \quad \dots \quad Av_n].$$

The right side equals

$$[\lambda_1 v_1 \quad \lambda_2 v_2 \quad \dots \quad \lambda_n v_n]$$

- Thus equating the columns of the matrix equation (2) gives

$$Av_i = \lambda_i v_i, \quad i = 1, \dots, n.$$

- This is independent of any assumptions on the eigenvectors. However, if the eigenvectors

are linearly independent then we can multiply with P^{-1} on both sides of (2) and get the similarity transform (1).

- This establishes facts **1.**, and **2.**.
- Numerical examples are tedious unless they are very small. Examples 3 and 4 in the text-book consider two 3×3 matrices.
- For a much simpler example, let's diagonalize the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$P = [v(2) \quad v(1)] \leftarrow$$

$$Av = 2v$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix} \Rightarrow \begin{bmatrix} x+y \\ 2y \end{bmatrix} \quad \begin{matrix} 2x = x+y \\ x = y \end{matrix}$$

$$x = x + y$$

$$y = 0 \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$$v(2) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$v(1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad P^{-1} = \frac{1}{-1} \begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \checkmark$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} P^{-1} A P$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

- What about the matrix

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad y=0$$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

- What about the matrix

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x+y \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \quad y = 0$$

- A matrix that is not diagonalizable is called **defective**.



Invertibility is unrelated to Diagonalizability.

- Complete this table with the simplest examples you can think of:

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

- Example:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a \\ b+c \\ c+d \\ d \end{bmatrix}$$

Loose Ends

- The following trick is sometimes useful. 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

$$A = PDP^{-1}. \quad (3)$$

- Using (3) also gives us some insight into what we are doing in terms of the eigenvectors and eigenvalues when multiplying a vector \mathbf{x} with A :

$$A\mathbf{x} = PDP^{-1}\mathbf{x}$$

We compose three linear functions. The first (multiplying with D^{-1}) converts \mathbf{x} into the coordinate vector of \mathbf{x} with respect to the eigenvector basis. The second (multiplying with D) modifies each eigenvector component individually. The third (multiplying with P) converts the coordinate vector with respect to the eigenvector basis back to the coordinate vector with respect to the standard basis.