Lecture 17

Monday, February 17, 2020

Recall the procedure to check if a linear map $f: V \to V$ is diagonalizable.

- Compute all the distinct eigenvalues of f. Label them as $\lambda_1, \lambda_2, \dots, \lambda_k$.
- To each eigenvalue λ_j, find a basis of the corresponding eigenspace E_{λ_j}.
- If $\dim E_{\lambda_1} + \dim E_{\lambda_2} + \dots + \dim E_{\lambda_k} = \dim V$ then f is diagonalizable.
  If $\dim E_{\lambda_1} + \dots + \dim E_{\lambda_k} < \dim V$ then f is not diagonalizable

In case f is diagonalizable, how do we find a basis $\mathcal{B}$ such that $[f]_{\mathcal{B}}$ is a diagonal matrix? Here is the procedure:

- We find a basis for each of the eigenspaces $E_{\lambda_1}, E_{\lambda_2}, \dots, E_{\lambda_k}$.
  Let us call B_1 a basis of E_{λ_1},
  B_2 a basis of E_{λ_2},
  ....
  B_k a basis of E_{λ_k}.
- Then $\mathcal{B} = B_1 \sqcup B_2 \sqcup \dots \sqcup B_k$.

What is $[f]_{\mathcal{B}}$?

Notice that B_1 is a subset of E_{λ_1}, so every vector in B_1 gets scaled by factor λ_1 under f. Likewise, every vector in B_2 gets scaled by factor λ_2 under f. And so on. The matrix that represents f in basis $\mathcal{B}$ is

$$[f]_{\mathcal{B}} = \begin{bmatrix} \boxed{\begin{matrix} \lambda_1 & & 0 \\ & \lambda_1 \ddots & \\ 0 & & \lambda_1 \end{matrix}} & & & 0 \\ & \boxed{\begin{matrix} \lambda_2 & 0 \\ 0 \ddots & \lambda_2 \end{matrix}} & & \\ & & \ddots & \\ 0 & & & \boxed{\begin{matrix} \lambda_k \ddots & 0 \\ 0 & \lambda_k \end{matrix}} \end{bmatrix}$$

The size of the first block is the size of $\mathcal{B}_1$. The size of the second block is the size of $\mathcal{B}_2$. And so on.

Ex: Consider the linear map $f: M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ given by

$$f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

We have seen in Lecture 14 that f has two distinct eigenvalues, namely $\lambda_1 = -1$ and $\lambda_2 = 1$. The corresponding spaces are

$$E_{-1} = \text{span}\left\{ \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}}_{v_1}, \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{v_2}, \underbrace{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}}_{v_3} \right\}$$

$$E_1 = \text{span}\left\{ \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{v_4} \right\}$$

A basis of E_{-1} is $\mathcal{B}_1 = \{v_1, v_2, v_3\}$

A basis of E_1 is $\mathcal{B}_2 = \{v_4\}$.

Then $\mathcal{B} = \mathcal{B}_1 \sqcup \mathcal{B}_2 = \{v_1, v_2, v_3, v_4\}$.

Let us check if f is indeed represented by a diagonal matrix. Because $v_1 \in E_{-1}$, it is scaled by factor -1 under f. That is,

$$f(v_1) = -v_1.$$

Similarly,

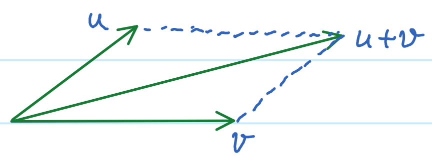
$$f(v_2) = -v_2$$
$$f(v_3) = -v_3$$
$$f(v_4) = v_4$$

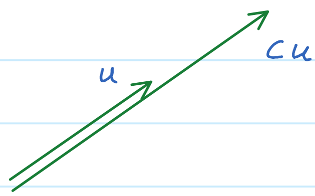
Thus,

$$[f]_{\mathcal{B}} = \begin{bmatrix} | & | & | & | \\ [f(v_1)]_{\mathcal{B}} & [f(v_2)]_{\mathcal{B}} & [f(v_3)]_{\mathcal{B}} & [f(v_4)]_{\mathcal{B}} \\ | & | & | & | \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & \boxed{1} \end{bmatrix}.$$

* Inner product:

A vector space is a set equipped with an addition rule and a scaling rule. One can add two vectors and scale a vector. The addition is called a binary operator because it takes in two vectors. The scaling is also a binary operator because it takes in a vector and a scaling factor.





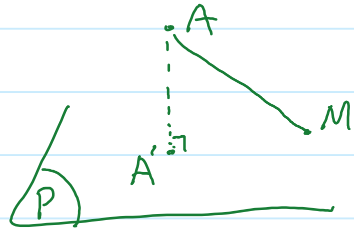
In $\mathbb{R}^2$ or $\mathbb{R}^3$, we can measure the angle between two vectors: the angle θ between vector a and vector b satisfies

$$\cos\theta = \frac{a \cdot b}{\|a\| \, \|b\|}.$$

We see the involvement of the dot product and the length of vectors. Length can be derived from the dot product because

$$\|a\| = \sqrt{a \cdot a}$$

It is natural to ask if one can define angles in an abstract vector spaces. For example, what is the angles between two matrices or between two polynomials? Having the concept of angle will give us more flexibility to work with abstract vectors. For example, we can say whether two abstract vectors (matrices, polynomials, ..) are perpendicular. This has applications in optimization problems. An example is that the distance from a point A outside of a plane (P) to a point M on the plane is minimum when M is the perpendicular projection of A on (P).



The addition and scaling are not sufficient to define angles between two vectors. We need another structure, something similar to the dot product on $\mathbb{R}^n$. This leads to the concept of inner product.

Definition: Let V be a vector space over a field $F = \mathbb{Q}, \mathbb{R}, \mathbb{C}$. A binary operator $(\cdot,\cdot) : V \times V \to F$ is said to be an inner product on V if it satisfies three following axioms:

(1) Linearity on the first argument:

$$(u+v, w) = (u,w) + (v,w) \quad \forall u,v,w \in V,$$
$$(cu, w) = c(u,w) \quad \forall u,w \in V, \ \forall c \in F.$$

(2) Conjugate symmetry:

$$(u,v) = \overline{(v,u)} \quad \forall u,v \in V.$$

Here $\bar{z}$ denotes the complex conjugate of a complex number z

For example, $\overline{1+2i} = 1-2i$, $\overline{-1} = -1$, $\bar{i} = -i$, ...

(3) Positive definiteness:

- $(u,u) \geq 0 \quad \forall u \in V.$
- If $(u,u) = 0$ then $u = 0$.

This definition mirrors the dot product in $\mathbb{R}^n$. In other words, inner product is a natural generalization of the dot product to an abstract vector.

Ex: On $\mathbb{R}^2$, let us define a binary operator $(\cdot,\cdot)$ as follows:

$$\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right) = x_1 y_1 + x_2 y_2$$

We can check that this operator satisfies all the axioms of inner product. Thus, it is an inner product on $\mathbb{R}^2$. [This is in fact the dot product on $\mathbb{R}^2$.]

* Note: we write a vector in $\mathbb{R}^2$ as $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ instead of the regular way (x_1, x_2) to avoid confusion. The notation $(\cdot,\cdot)$ is reserved for inner product.

Ex: On $\mathbb{R}^2$, consider a binary operator $(\cdot,\cdot)$ given by

$$\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right) = x_1 + y_1$$

This operator is not an inner product because

$$\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) = 0 + 0 = 0$$

but $\begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. In other words, the positive definiteness is violated.