Discussion 2A/2B Notes

Week 1 : Jan. 10, 12

Notations

- V : a vector space, usually $\mathbb{R}^d$ with $d=1,2,3,\cdots$
  (over the field $\mathbb{R}$, unless specified)
- For a set S, $|S|$ denotes the number of elements in S.
- For a set S, S^m denotes the Cartesian product

$$S^m = \underbrace{S \times S \times \cdots \times S}_{m\text{-times}} := \{(s_1, s_2, \cdots, s_m) : s_i \in S \text{ for } i = 1, \cdots, m\}$$

- For sets $A, B \subseteq V$, $A \setminus B$ denotes $A \cap B^c$.
  In particular, $V \setminus \{0\}$ denotes the set of all nonzero vectors of V.

Linear Independence

Def. Let $S = \{v_1, \cdots, v_m\} \subseteq V$.

If $\exists (a_1, a_2, \cdots, a_m) \in \mathbb{R}^m \setminus \{0\}$ s.t. $\sum_{i=1}^m a_i v_i = 0$, S is said to be linearly dependent

- S is said to be linearly independent iff it's not linearly dependent.

(That is, $\sum_{i=1}^m a_i v_i = 0$ admits only the trivial solution $a_1 = a_2 = \cdots = a_m = 0$.)

Ex.
- If $\dim V = d < \infty$ and $m = |S| > d$, then S is always linearly dependent.

- In other words, if S is linearly independent, then $|S| \leq d$.

- In fact, the above bound is tight. i.e., $\exists S \subseteq V$, lin. indep., with $|S| = \dim V$.

  e.g. When $V = \mathbb{R}^d$, choose $S = \{e_1, e_2, \cdots, e_d\}$ where

$$(e_i)_j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \quad \text{for all } 1 \leq i, j \leq d.$$

  (These sequence of vectors $(e_1, \cdots, e_d)$ is called the standard basis of $\mathbb{R}^d$.)

- In such a case, i.e., $|S| = \dim V$ and S is lin. indep., then

  * S spans V.

    i.e., $\operatorname{span} S := \left\{ \sum_{i=1}^d a_i v_i : (a_1, \cdots, a_d) \in \mathbb{R}^d, (v_1, \cdots, v_d) \in S^d \right\} = V.$

  * Moreover, each vector in V has a unique representation as a linear combination of S.

    i.e., $\forall v \in V$, $\exists! (a_1, \cdots, a_d) \in \mathbb{R}^d$ s.t. $v = \sum_{i=1}^d a_i v_i$.

ⓞ Eigenvalues and Eigenvectors

- Let $A : V \to V$ be linear

  (i.e., $\forall x, x' \in V$ and $\forall c, c' \in \mathbb{R}$, $A(cx + c'x') = cAx + c'Ax'$.)

non-zero vectors only!

Def If a pair $(\lambda, v) \in \mathbb{F} \times (V \setminus \{0\})$ satisfies $Av = \lambda v$,
v is said to be an eigenvector of A, and
λ is said to be an eigenvalue of A (associated with v.)

Understanding the meaning of eigenvectors / eigenvalues is extremely important (beyond the simple definition "$Av = \lambda v$".)

Let $(\lambda_1, v_1), \cdots, (\lambda_d, v_d)$ are eigenvalue-eigenvector pairs for A.

If $v_1, \cdots, v_d$ are linearly independent, then $\{v_1, \cdots, v_d\}$ spans $\mathbb{R}^d$, and thus

$\forall x \in \mathbb{R}^d \ \exists \ (a_1, \cdots, a_d) \in \mathbb{R}^d$ s.t. $x = \sum_{i=1}^{d} a_i v_i$.

Then,
$$
\begin{aligned}
Ax &= A\left(\sum_{i=1}^{d} a_i v_i\right) \\
&= \sum_{i=1}^{d} a_i (Av_i) && (\textit{linearity}) \\
&= \sum_{i=1}^{d} a_i \lambda_i v_i && (\textit{eigenvectors})
\end{aligned}
$$

and this is a simple linear combination of v_i's, instead of a (possibly) complicated map A.

Thus, applying A to x can be understood as

Step ① finding $(a_1, \cdots, a_d)$ s.t. $x = \sum_{i=1}^{d} a_i v_i$ — (*)

Step ② Recovering a lin. comb. $\sum_{i=1}^{d} a_i \lambda_i v_i$.

In general step ① is more difficult. However, it is still straightforward, because (*) means

$$x = Pa, \quad \text{where} \quad P = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_d \\ | & | & & | \end{bmatrix}_{d \times d}, \quad a = \begin{bmatrix} a_1 \\ \vdots \\ a_d \end{bmatrix} \in \mathbb{R}^d.$$

Therefore, a is simply $P^{-1}x$. Then Step ② becomes applying the following map to the coefficients:

$$D : \begin{bmatrix} a_1 \\ \vdots \\ a_d \end{bmatrix} \longmapsto \begin{bmatrix} \lambda_1 a_1 \\ \vdots \\ \lambda_d a_d \end{bmatrix}.$$

Fortunately, this linear transformation can be represented by a very simple matrix :

$$D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_d \end{bmatrix}.$$

Finally, recovering a linear combination is just multiplying P again because

$$P : \begin{bmatrix} c_1 \\ \vdots \\ c_d \end{bmatrix} \mapsto \sum_{i=1}^{d} c_i v_i$$

Thus, $Ax = PDP^{-1}x \quad \forall x \in \mathbb{R}^d$, i.e., $A = PDP^{-1}$ and

it's called a diagonalization of A.

Note • If A is real and symmetric, i.e., $A \in \mathbb{R}^{d \times d}$ and $A^T = A$, then

A is diagonalizable by orthogonal matrices. $\Leftrightarrow$ P is orthogonal $\Leftrightarrow$ $P^{-1} = P^T$

$$\Leftrightarrow \langle v_i, v_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

• Remember that we assumed that $\{v_1, \cdots, v_d\}$ is linearly independent.

Otherwise, we can't find a diagonalization of A and things can get complicated.

However, the Jordan normal form of A still always exists:

$A = PJP^{-1}$ and $J =$ e.g. $\begin{bmatrix} \boxed{\begin{matrix} \lambda_1 & 1 & \\ & \lambda_1 & 1 \\ & & \lambda_1 \end{matrix}} & & \\ & \boxed{\begin{matrix} \lambda_2 & 1 \\ & \lambda_2 \end{matrix}} & \\ & & \ddots \end{bmatrix}$ a block diagonal matrix, consists of "Jordan blocks".

and $P = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_d \\ | & & | \end{bmatrix}$ where v_i's are generalized eigenvectors.

Linear ODE's

First recall that in 1-D, the general solution to a linear ODE

$$f'(x) = c f(x)$$

is $f(x) = A e^{cx}$, where A is a constant determined by other conditions (e.g. initial condition $f(0)$: $f(x) = f(0) e^{cx}$ — $(*)$)

An analogy can be made in $\mathbb{R}^d$, too. Consider

$$(**) \quad \dot{x}(t) = A\, x(t)$$

where $x : \mathbb{R} \to \mathbb{R}^d$, $t \mapsto x(t)$ and $A \in \mathbb{R}^{d \times d}$ is a constant matrix.

If A is diagonalizable, $A = PDP^{-1}$ and $(**)$ can be re-written with the change of variables $y(t) := P^{-1} x(t)$:

$$(***) \quad \dot{y} = D y$$

Unlike $(**)$, $(***)$ is very easy to solve because the system of ODE's (i.e., y_i's) is now decoupled:

$$(***) \iff \begin{cases} \dot{y}_1 = \lambda_1 y_1 \\ \dot{y}_2 = \lambda_2 y_2 \\ \quad \vdots \\ \dot{y}_d = \lambda_d y_d \end{cases} \iff y_i = y_i(0) e^{\lambda_i t} \iff y(t) = \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_d t} \end{bmatrix} y(0)$$

Now, we define the matrix exponential for diagonal matrices $\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{bmatrix} \mapsto \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_d t} \end{bmatrix}$ and write $y(t) = e^D y(0)$.

Finally, recover $x(t)$ by $Py(t) \Rightarrow x(t) = P e^D P x(0)$.

If we define the matrix exponential by Taylor series: $e^A := \sum_{n=0}^{\infty} \frac{A^n}{n!}$ for A where the series converges, then if $A = PDP^{-1}$, $A^n = PD^nP^{-1}$ and we have $\sum_{n=0}^{\infty} \frac{A^n}{n!} = P \sum_{n=0}^{\infty} \frac{D^n}{n!} P^{-1}$ $= Pe^DP^{-1}$, and this coincides with the def. of matrix exp. for diagonal matrices.

Moreover, we can also write $x(t) = e^A x(0)$, which generalizes $(*)$. $\square$