MATH-GA2120 Linear Algebra II

Diagonalizable Linear Maps
Inner Product on $\mathbb{F}^n$
Inner Product on Abstract Vector Space
Cauchy-Schwarz Inequality

Deane Yang

Courant Institute of Mathematical Sciences
New York University

February 15, 2024

Diagonal Linear Transformation

- Let $\dim V = n$
- Let $L : V \to V$ be a linear transformation
- Suppose L has n linearly independent eigenvectors $e_1, \ldots, e_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$
- Then with respect to the basis $E = (e_1, \ldots, e_n)$,

$$L(e_k) = e_k \lambda_k$$

- Equivalently,

$$\begin{bmatrix} L(e_1) & \cdots & L(e_n) \end{bmatrix} = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Diagonal Linear Transformation

- Conversely, suppose $L : V \to V$ is a linear transformation and E is a basis such that

$$L(E) = ED,$$

  where D is a diagonal matrix
- Then

$$L(e_k) = e_j D^j_k = e_k D^k_k$$

- Therefore, L has eigenvalues $D^1_1, \ldots, D^n_n$ with eigenvectors $e_1, \ldots, e_n$ respectively

Diagonalizable Linear Transformation

- Let $L : V \to V$ be a diagonal linear transformation
- If E is a basis of eigenvectors, then

$$L(E) = ED,$$

  where D is a diagonal matrix
- Given any basis F, there is an invertible matrix M such that

$$F = EM$$

  and vice versa
- There is a matrix A such that

$$L(F) = FA$$

- Therefore,

$$ED = L(E) = L(FM^{-1}) = L(F)M^{-1} = FAM^{-1} = EMAM^{-1}$$

- I.e., M and D are similar

Diagonalizable Linear Transformation and Matrix

- A linear transformation $L : V \to V$ is **diagonalizable** if any of the following equivalent conditions hold:
  - There exists a basis of V consisting of eigenvectors
  - There exists a basis E such that $L(E) = ED$, where D is a diagonal matrix
  - Given any basis F and matrix A such that

    $$L(F) = FA,$$

    A is similar to a diagonal matrix
  - A matrix A is **diagonalizable** if it is similar to a diagonal matrix

Linear Transformation With Distinct Eigenvalues

- Let $\dim(V) = n$ and $L : V \to V$ be a linear transformation with n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, i.e.,

$$j \neq k \implies \lambda_j \neq \lambda_k$$

- Let $v_1, \ldots, v_n$ be eigenvectors of $\lambda_1, \ldots, \lambda_n$ respectively
- Suppose $v_1, \ldots, v_{k-1}$ are linearly independent
- If $a^1 v_1 + \cdots + a^k v_k w = 0$, then

$$\begin{aligned}
0 &= (L - \lambda_k I)(a^1 v_1 + \cdots + a^k v_k) \\
&= a^1(Lv_1) - \lambda_k v_1) + \cdots + a^k(L(v_k) - \lambda_k v_k) \\
&= a^1(\lambda_1 - \lambda_k)v_1 + \cdots + a^k(\lambda_k - \lambda_k)v_k \\
&= a^1(\lambda_1 - \lambda_k)v_1 + \cdots + a^{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1},
\end{aligned}$$

- Therefore, $a^1(\lambda_1 - \lambda_k) = \cdots = a^{k-1}(\lambda_{k-1} - \lambda_k) = 0$

Linear Transformation With Distinct Eigenvalues

- Since $v_1, \dots, v_{k-1}$ are linearly independent, it follows that

$$a^1(\lambda_1 - \lambda_k) = \cdots = a^{k-1}(\lambda_{k-1} - \lambda_k) = 0$$

- Since the eigenvalues are distinct, this implies that

$$a^1 = \cdots = a^{k-1} = 0$$

- By assumption, $a^1 v_1 + \cdots + a^k v_k w = 0$ and therefore $a^k = 0$
- It follows by induction that $v_1, \dots, v_n$ form a basis of V
- Therefore, L is diagonalizable
- **Conclusion:** Any linear transformation with n distinct eigenvalues is diagonalizable

Direct Sum of Subspaces

- Let $V_1, \ldots, V_k$ be subspaces of V
- $\{V_1, \ldots, V_k\}$ is a **linearly independent** set of subspaces if for any nonzero vectors

$$v_1 \in V_1,\ v_2 \in V_2, \ldots, v_k \in V_k$$

  are linearly independent
- Equivalently, $\{V_1, \ldots, V_k\}$ is linearly independent if for any $v_1 \in V_1, \ldots, v_k \in V_k$,

$$v_1 + v_2 + \cdots + v_k = 0 \implies v_1 = v_2 = \cdots = v_k$$

- Equivalently, $\{V_1, \ldots, V_k\}$ is linearly independent if for any $v_1, w_1 \in V_1, \ldots, v_k, w_k \in V_k$,

$$v_1 + v_2 + \cdots + v_k = w_1 + w_2 + \cdots + w_k \implies v_1 = w_1, \ldots, v_k = w_k$$

- If $\{V_1, V_2, \ldots, V_k\}$ is linearly independent, then their **direct sum** is defined to be

$$V_1 \oplus V_2 \oplus \cdots \oplus V_k = \text{span}(V_1 \cup V_2 \cup \cdots \cup V_k)$$

Examples

- $\{S_1, S_2\}$, where $S_1, S_2 \subset \mathbb{F}^3$ are given by
$$S_1 = \text{span}(e_1)$$
$$S_2 = \text{span}(e_2),$$
is linearly independent
- If $\{v_1, \ldots, v_k\}$ is linearly independent and
$$\forall 1 \leq j \leq k, \;\; V_j = \text{span}(v_j),$$
then $\{V_1, \ldots, V_k\}$ is a linearly independent set of subspaces
- If (e_1, e_2, e_3, e_4) is a basis of V and
$$S = \text{span}(e_1, e_2, e_3), \;\; T = \text{span}(e_4),$$
then $V = S \oplus T$

Eigenspaces of Distinct Eigenvalues are Linearly Independent (Part 1)

- If $\lambda_1, \ldots, \lambda_k$ are distinct eigenvalues of $L : V \to V$, then their eigenspaces $E_{\lambda_1}, \ldots, E_{\lambda_k}$ are linearly independent
- Prove by induction that for any $1 \leq j \leq k$,

$$v_1 + \cdots + v_j = 0 \implies v_1 = \cdots = v_j = 0$$

- This holds for $j = 1$
- Inductive step: Assume that it holds for $1 \leq j < k$ and prove it holds for $j + 1$

Eigenspaces of Distinct Eigenvalues are Linearly Independent (Part 2)

- Suppose $v_1 \in E_{\lambda_1}, \ldots, v_{j+1} \in E_{\lambda_{j+1}}$ satisfy

$$v_1 + \cdots + v_{j+1} = 0 \tag{1}$$

- It follows that

$$\begin{aligned} 0 &= (L - \lambda_{j+1} I)(v_1 + \cdots + v_{j+1}) \\ &= (\lambda_1 - \lambda_{j+1}) v_1 + \cdots + (\lambda_j - \lambda_{j+1}) v_j \end{aligned}$$

- By the inductive assumption,

$$(\lambda_1 - \lambda_{j+1}) v_1 = \cdots = (\lambda_j - \lambda_{j+1}) v_j = 0$$

- Since $\lambda_i - \lambda_{j+1} \neq 0$ for each $1 \leq i \leq j$,

$$v_1 = \cdots = v_j = 0$$

- By (1), it follows that $v_{j+1} = 0$

Eigenspaces of Distinct Eigenvalues are Linearly Independent (Part 3)

- By induction,

$$v_1 + \cdots + v_k = 0 \implies v_1 = \cdots = v_k = 0$$

- This implies that $E_{\lambda_1}, \ldots, E_{\lambda_k}$ are linearly independent

Diagonalizability of a Linear Transformation (Part 1)

- Let $\lambda_1, \ldots, \lambda_k$ be the eigenvalues of $L : V \to V$
- L is diagonalizable if and only if

$$\dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}) = \dim V$$

- Let $n_0 = 0$ and, for $1 \leq j \leq k$, let

$$n_j = \dim(E_{\lambda_j})$$
$$N_j = n_1 + \cdots + n_j$$

- For each $1 \leq j \leq k$, let

$$(v_{N_{j-1}+1}, \cdots, v_{N_j})$$

be a basis of E_{λ_j}

Diagonalizability of a Linear Transformation (Part 2)

- Suppose

$$a^1 v_1 + \cdots + a^n v_n = 0,$$

- For each $1 \leq j \leq k$, let

$$w_j = a^{N_{j-1}+1} v_{N_{j-1}} + \cdots + a^{N_j} v_{N_j} \in E_{\lambda_j}$$

- Since $w_1 + \cdots + w_k = 0$, it follows that

$$w_1 = \cdots = w_k = 0$$

- For each $1 \leq j \leq k$,

$$0 = w_j = a^{N_{j-1}+1} v_{N_{j-1}} + \cdots + a^{N_j} v_{N_j},$$

  which implies $a^{N_{j-1}+1} = \cdots = a^{N_j} = 0$
- Therefore, $(v_1, \ldots, v_n)$ is a basis of V
- L is diagonal with respect to this basis

Dot Product on $\mathbb{R}^n$

- Recall that the **dot product** of

$$v = \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix}, \ w = \begin{bmatrix} w^1 \\ \vdots \\ w^n \end{bmatrix} \in \mathbb{R}^n$$

  is defined to be

$$v \cdot w = v^1 w^1 + \cdots + v^n w^n = v^T w = w^T v$$

- The **norm** or **magnitude** of $v \in \mathbb{R}^n$ is defined to be

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

- If v and w are nonzero and the angle at 0 from v to w is θ, then

$$\cos\theta = \frac{v \cdot w}{|v||w|}$$

Properties of Dot Product

- The dot product is **bilinear** because for any $a, b \in \mathbb{R}$ and $u, v, w \in \mathbb{R}^n$,

$$(au + bv) \cdot w = a(u \cdot w) + b(v \cdot w)$$
$$u \cdot (av + bw) = a(u \cdot v) + b(u \cdot w)$$

- It is **symmetric**, because for any $v, w \in \mathbb{R}^n$,

$$v \cdot w = w \cdot v$$

- It is **positive definite**, because for any $v \in \mathbb{R}^n$,

$$v \cdot v \geq 0$$

  and

$$v \cdot v > 0 \iff v \neq 0$$

Inner Product on Real Vector Space

- Let V be an n-dimensional real vector space
- Consider a function

$$\alpha : V \times V \to \mathbb{R}$$

- It is **bilinear** if for any $a, b \in \mathbb{R}$ and $u, v, w \in \mathbb{R}^n$,

$$\alpha(au + bv, w) = a\alpha(u, w) + b\alpha(v, w)$$
$$\alpha(u, av + bw) = a\alpha(u, v) + b\alpha(u, w)$$

- It is **symmetric** if for any $v, w \in \mathbb{R}^n$,

$$\alpha(v, w) = \alpha(w, v)$$

- It is **positive definite** if for any $v \in \mathbb{R}^n$,

$$\alpha(v, v) \geq 0$$

  and

$$\alpha(v, v) > 0 \iff v \neq 0$$

- Any positive definite symmetric bilinear function on a **real** vector space V is called an **inner product**

Hermitian Inner Product on $\mathbb{C}^n$

- Recall that if $z = x + iy \in \mathbb{C}$, then
$$\bar{z} = x - iy \text{ and } z\bar{z} = \bar{z}z = x^2 + y^2$$
- If A is a complex matrix, its **Hermitian adjoint** is defined to be
$$A^* = \bar{A}^*$$
- The Hermitian inner product on $\mathbb{C}^n$ of
$$v = \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix}, \; w = \begin{bmatrix} w^1 \\ \vdots \\ w^n \end{bmatrix} \in \mathbb{C}^n$$
is defined to be
$$(v, w) = v^1\bar{w}^1 + \cdots + v^n\bar{w}^n = v^T\bar{w} = \bar{w}^Tv = w^*v \in \mathbb{C},$$
- The **norm** of $v \in \mathbb{C}^n$ is defined to be
$$|v| = \|v\| = \sqrt{(v, v)}$$
- **No** geometric interpretation of the Hermitian inner product

Not a Real Inner Product

- **Not** bilinear, because if $c \in \mathbb{C}$,

$$(v, cw) = \bar{c}(v, w)$$

- **Not** symmetric, because

$$(w, v) = \overline{(v, w)}$$

- It is positive definite, because for any $v \in \mathbb{C}^n$, $(v, v) \in \mathbb{R}$,

$$(v, v) = v^1 \bar{v}^1 + \cdots + v^n \bar{v}^n = |v^1|^2 + \cdots + |v^n|^2 \geq 0,$$

and

$$(v, v) \neq 0 \iff v \neq 0$$

Properties of Hermitian Inner Product on $\mathbb{C}^n$

- It is a linear function of the first argument, because for any $a, b \in \mathbb{C}$, $u, v, w \in \mathbb{C}^n$,

$$(au + bv, w) = a(u, w) + b(v, w)$$

- It is **Hermitian**, which means

$$(v, w) = \overline{(w, v)}$$

- Therefore, for any $a, b \in \mathbb{C}$ and $u, v, w \in \mathbb{C}^n$,

$$(u, av + bw) = \bar{a}(u, v) + \bar{b}(a, w)$$

Inner Product of a Vector Space Over $\mathbb{F}$

- Assume $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$
- An **inner product** over a vector space V is a function

  $$(\cdot, \cdot) : V \times V \to \mathbb{F}$$

  with the following properties: For any $a, b \in F$ and $u, v, w \in V$,

  $$\begin{aligned} (au + bv, w) &= a(u, w) + b(v, w) \\ (w, v) &= \overline{(v, w)} \\ (v, v) &\geq 0 \\ (v, v) \neq 0 &\iff v \neq 0 \end{aligned}$$

- If $\mathbb{F} = \mathbb{R}$, this is the same definition as before
- If $\mathbb{F} = \mathbb{C}$, this is the definition of a Hermitian inner product

Examples

- For each $v \in \mathbb{F}^n$, denote $v^* = \bar{v}^T$
- The standard inner product on $\mathbb{F}^n$ is

$$(v, w) = w^* v,$$

  which is the dot product on $\mathbb{R}^n$ and the standard Hermitian inner product on $\mathbb{C}^n$
- An inner product on the space of polynomials of degree n or less and with coefficients in $\mathbb{F}$ is

$$(f, g) = \int_{t=0}^{t=1} f(t)\overline{g(t)}\, dt$$

- An inner product on the space of matrices with n rows and m columns is

$$(A, B) = \operatorname{trace}(B^* A) = \sum_{1 \le k \le m} \sum_{1 \le j \le n} \bar{B}^j_k A^j_k,$$

  where $B^* = \bar{B}^T$

Nondegeneracy Property

- Fact: If a vector $v \in V$ satisfies the following property:

$$\forall w \in V,\ (v, w) = 0,$$

  then $v = 0$
  - Proof: Setting $w = v$, it follows that

$$(v, v) = 0 \text{ and therefore } v = 0$$

- Corollary: If $v_1, v_2 \in V$ satisfy the property that

$$\forall w \in V,\ (v_1, w) = (v_2, w),$$

  then $v_1 = v_2$
- Corollary: If '$L_1, L_2 : V \to W$ are linear maps such that

$$\forall v \in V,\ w \in W,\ (L_1(v), w) = (L_2(v), w),$$

  then $L_1 = L_2$
  - Proof: Given $v \in V$,

$$\forall w \in W,\ (L_1(v), w) = (L_2(v), w),$$

    which implies $L_1(v) = L_2(v)$
  - Since this holds for all $v \in V$, it follows that $L_1 = L_2$

Fundamental Inequalities

- **Cauchy-Schwarz inequality:** For any $v, w \in V$,

  $$|(v, w)| \leq |v||w|$$

  and

  $$|(v, w)| = |v||w|$$

  if and only if there exists $s \in \mathbb{F}$ such that

  $$v = sw \text{ or } w = sv$$

- **Triangle inequality:** For any $v, w \in V$,

  $$|v + w| \leq |v| + |w|$$

  and

  $$|v + w| = |v| + |w|$$

  if and only if $v = \pm w$

Proof When $\mathbb{F} = \mathbb{R}$

- If $v = 0$ or $w = 0$, equality holds
- Let

$$\begin{aligned} f(t) &= |v - tw|^2 \\ &= (v - tw, v - tw) \\ &= |v|^2 - 2t(v, w) + t^2|w|^2 \\ &= \left(t|w| - \frac{(v, w)}{|w|} \right)^2 + |v|^2 - \frac{(v, w)^2}{|w|^2} \end{aligned}$$

- f has a unique minimum when $t = t_{\min}$, where

$$t_{\min} = \frac{(v, w)}{|w|} \text{ and } f(t_{\min}) = |v|^2 - \frac{(v, w)^2}{|w|^2}$$

Proof of Cauchy-Schwarz (Part 1)

- If $v = 0$ or $w = 0$, equality holds
- If $w \neq 0$, let $f : \mathbb{F} \to \mathbb{R}$ be the function

$$\begin{aligned} f(t) &= |v - tw|^2 \\ &= (v - tw, v - tw) \\ &= |v|^2 - t(w, v) - \bar{t}(v, w) + |t|^2|w|^2 \end{aligned}$$

- If f has a minimum at $t_0 \in \mathbb{F}$, then its directional derivative at t_0 is zero in any direction $\dot{t}$

$$\begin{aligned} 0 &= \left.\frac{d}{ds}\right|_{s=0} f(t_0 + s\dot{t}) \\ &= -\dot{t}(w, v) - \bar{\dot{t}}(v, w) + (t_0\bar{\dot{t}} + \bar{t}_0\dot{t})|w|^2 \\ &= \dot{t}(\bar{t}_0 - (w, v)) + \bar{\dot{t}}(t_0|w|^2 - (v, w)) \\ &= \dot{t}(\overline{t_0 - (v, w)}) + \bar{\dot{t}}(t_0|w|^2 - (v, w)) \end{aligned}$$

Proof of Cauchy-Schwarz (Part 2)

- In particular, if

$$\dot{t} = t_0|w|^2 - (v, w),$$

we get

$$|t_0|w|^2 - (v, w)|^2 = 0,$$

- Therefore, the only critical point of f is

$$t_0 = \frac{(v, w)}{|w|^2}$$

- Since f is always nonnegative, it follows that

$$0 \leq f(t_0) = |v|^2 - \frac{|(v, w)|^2}{|w|^2}$$

which implies the Cauchy-Schwarz inequality

Proof of Cauchy-Schwarz (Part 3)

- If $w \neq 0$ and $|(v, w)| = |v||w|$, then

$$0 = |v|^2 - \frac{|(v, w)|^2}{|w|^2} = f(t_0) = |v - t_0 w|^2,$$

which implies that

$$v = t_0 w$$