TTIC 31150/CMSC 31150 Mathematical Toolkit (Spring 2023)

Avrim Blum and Ali Vakilian

Lecture 3: Eigenvectors and Eigenvalues, Inner Products

Recap

L1
- Vector spaces, linear dependence / independence, span, basis, Steinitz Exchange Principle, dimension of vector space

L2
- Lagrange interpolation
- Secret sharing
- Linear transformations
  - LT uniquely determined by action on a basis. Connection to matrices.
  - Kernel (nullspace) and image. Rank-nullity theorem.

Definition of Eigenvectors and Eigenvalues

Definition 1.1 *Let V be a vector space over the field $\mathbb{F}$ and let $\varphi : V \to V$ be a linear transformation. $\lambda \in \mathbb{F}$ is said to be an* eigenvalue *of φ if there exists $v \in V \setminus \{0_V\}$ such that $\varphi(v) = \lambda \cdot v$. Such a vector v is called an* eigenvector *corresponding to the eigenvalue λ. The set of eigenvalues of φ is called its* spectrum*:*

$$\mathrm{spec}(\varphi) = \{\lambda \mid \lambda \textit{ is an eigenvalue of } \varphi\} .$$

Simple example: what are eigenvectors and eigenvalues of φ_A for $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$?

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector of eigenvalue 2, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is an eigenvector of eigenvalue 3.

Definition of Eigenvectors and Eigenvalues

Definition 1.1 *Let V be a vector space over the field $\mathbb{F}$ and let $\varphi : V \to V$ be a linear transformation. $\lambda \in \mathbb{F}$ is said to be an* eigenvalue *of φ if there exists $v \in V \setminus \{0_V\}$ such that $\varphi(v) = \lambda \cdot v$. Such a vector v is called an* eigenvector *corresponding to the eigenvalue λ. The set of eigenvalues of φ is called its* spectrum*:*

$$\text{spec}(\varphi) = \{\lambda \mid \lambda \textit{ is an eigenvalue of } \varphi\} \ .$$

Another example: differentiation is a linear transformation on the class of infinitely-differentiable real-valued functions. Each function of the form $ce^{\lambda x}$ is an eigenvector of eigenvalue λ.

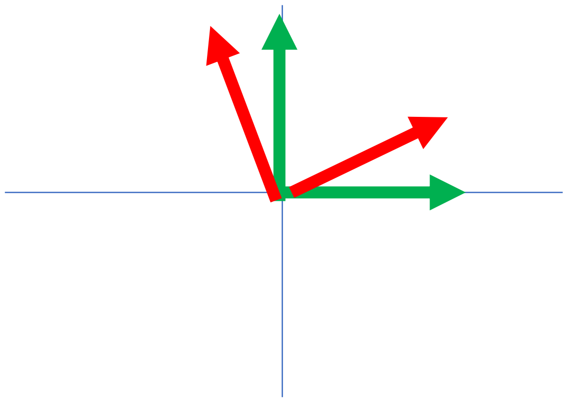
So, $\text{spec}(\varphi) = \mathbb{R}$.

Definition of Eigenvectors and Eigenvalues

Definition 1.1 *Let V be a vector space over the field $\mathbb{F}$ and let $\varphi : V \to V$ be a linear transformation. $\lambda \in \mathbb{F}$ is said to be an* eigenvalue *of φ if there exists $v \in V \setminus \{0_V\}$ such that $\varphi(v) = \lambda \cdot v$. Such a vector v is called an* eigenvector *corresponding to the eigenvalue λ. The set of eigenvalues of φ is called its* spectrum*:*

$$\text{spec}(\varphi) = \{\lambda \mid \lambda \textit{ is an eigenvalue of } \varphi\} .$$

Can also have no eigenvectors/eigenvalues, such as for a rotation matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.



$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \to \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix} \to \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$

Important Properties

Set of eigenvectors of eigenvalue λ

Proposition 1.4 *Let $U_\lambda = \{v \in V \mid \varphi(v) = \lambda \cdot v\}$. Then for each $\lambda \in \mathbb{F}$, U_λ is a subspace of V.*

Note that $U_\lambda = \{0_V\}$ if λ is not an eigenvalue. The dimension of this subspace is called the geometric multiplicity of the eigenvalue λ.

Why subspace?

If $v_1, v_2 \in U_\lambda$ then $a_1 v_1 + a_2 v_2 \in U_\lambda$, because

Same λ!

$$\varphi(a_1 v_1 + a_2 v_2) = a_1 \varphi(v_1) + a_2 \varphi(v_2) = \lambda(a_1 v_1 + a_2 v_2).$$

Important Properties

Proposition 1.5 *Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of φ with associated eigenvectors $v_1, \ldots, v_k$. Then the set $\{v_1, \ldots, v_k\}$ is linearly independent.*

So, eigenvectors of the same eigenvalue form a subspace, and eigenvectors with different eigenvalues are linearly independent.

Important Properties

Proposition 1.5 *Let $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues of φ with associated eigenvectors $v_1, \ldots, v_k$. Then the set $\{v_1, \ldots, v_k\}$ is linearly independent.*

Proof: We can prove by induction on k (base case of $k = 1$ is immediate). Assume true for $k - 1$ and suppose it was not true for k. Then one of the vectors, say v_k could be written as a linear combination of the others: $v_k = a_1 v_1 + \ldots + a_{k-1} v_{k-1}$ where the a_i are not all 0. Applying φ we get $\lambda_k v_k = \lambda_1 a_1 v_1 + \ldots + \lambda_{k-1} a_{k-1} v_{k-1}$. But now re-writing the left-hand-side in terms of $v_1, \ldots, v_{k-1}$ and re-grouping, we get $(\lambda_k - \lambda_1) a_1 v_1 + \ldots + (\lambda_k - \lambda_{k-1}) a_{k-1} v_{k-1} = 0$. Since the λ's are all distinct, this is a nonzero linear combination summing to 0, which contradicts our inductive assumption. ■

Definition 1.6 *A transformation $\varphi : V \to V$ is said to be* diagonalizable *if there exists a basis of V comprising of eigenvectors of φ.*

(E.g., if V has dimension n, and φ has n distinct eigenvalues)

(But distinctness not required. E.g., Identity transformation)

Inner Products

Definition 2.1 *Let V be a vector space over a field $\mathbb{F}$ (which is taken to be $\mathbb{R}$ or $\mathbb{C}$). A function $\mu : V \times V \to \mathbb{F}$ is an inner product if*

- *The function $\mu(u, \cdot) : V \to \mathbb{F}$ is a linear transformation for every $u \in V$. So, $\mu(u, cv) = c\mu(u,v)$ and $\mu(u, v+w) = \mu(u,v) + \mu(u,w)$.*
- *The function satisfies $\mu(u,v) = \overline{\mu(v,u)}$. (Complex conjugate)*
- *$\mu(v,v) \in \mathbb{R}_{\geq 0}$ for all $v \in V$ and is 0 only for $v = 0_V$. This is called positive semidefiniteness.*

We write the inner product corresponding to μ as $\langle u, v \rangle$. Define $\|v\| = \sqrt{\langle v, v \rangle}$

If over $\mathbb{R}$, then μ is symmetric, so $\mu(cu, v) = c\mu(u,v)$. Over $\mathbb{C}$, get $\mu(cu,v) = \bar{c}\mu(u,v)$.

(some definitions do it the other way around, with $\mu(\cdot, v)$ as a linear transformation)

Either way, you also get $\mu(u_1 + u_2, v) = \mu(u_1, v) + \mu(u_2, v)$.

Inner Products

Definition 2.1 *Let V be a vector space over a field $\mathbb{F}$ (which is taken to be $\mathbb{R}$ or $\mathbb{C}$). A function $\mu : V \times V \to \mathbb{F}$ is an inner product if*

- *The function $\mu(u, \cdot) : V \to \mathbb{F}$ is a linear transformation for every $u \in V$. So, $\mu(u, cv) = c\mu(u, v)$ and $\mu(u, v + w) = \mu(u, v) + \mu(u, w)$.*
- *The function satisfies $\mu(u, v) = \overline{\mu(v, u)}$. (Complex conjugate)*
- *$\mu(v, v) \in \mathbb{R}_{\geq 0}$ for all $v \in V$ and is 0 only for $v = 0_V$. This is called positive semidefiniteness.*

We write the inner product corresponding to μ as $\langle u, v \rangle$. Define $\|v\| = \sqrt{\langle v, v \rangle}$

If over $\mathbb{R}$, then μ is symmetric, so $\mu(cu, v) = c\mu(u, v)$. Over $\mathbb{C}$, get $\mu(cu, v) = \bar{c}\mu(u, v)$.

What about $\mu(cu, cv)$? Get $c\bar{c}\mu(u, v) = |c|^2\mu(u, v)$.

$$(a + bi)(a - bi) = a^2 + b^2$$

Inner Products Examples

- $\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx$ is an inner product over the vector space of continuous functions from [-1,1] to $\mathbb{R}$.

- For $x, y \in \mathbb{R}^2$, $\langle x, y \rangle = x_1 y_1 + x_2 y_2$ is the usual inner product. But $\langle x, y \rangle = 2x_1 y_1 + x_2 y_2 + x_1 y_2/2 + x_2 y_1/2$ also defines an inner product.

  - $\langle x, cy \rangle = c\langle x, y \rangle$.
  - For $y = u + v$, get $\langle x, y \rangle = \langle x, u \rangle + \langle x, v \rangle$.
  - $\langle x, x \rangle \geq 0$, since $2x_1^2 + x_2^2 + x_1 x_2 \geq \frac{1}{4}x_1^2 + x_2^2 + x_1 x_2 = \left(\frac{1}{2}x_1 + x_2\right)^2 \geq 0$.

Important Properties

Proposition 3.1 (Cauchy-Schwartz): for any two vectors $u, v \in V$,

$$|\langle u, v\rangle|^2 \leq \langle u, u\rangle \cdot \langle v, v\rangle$$

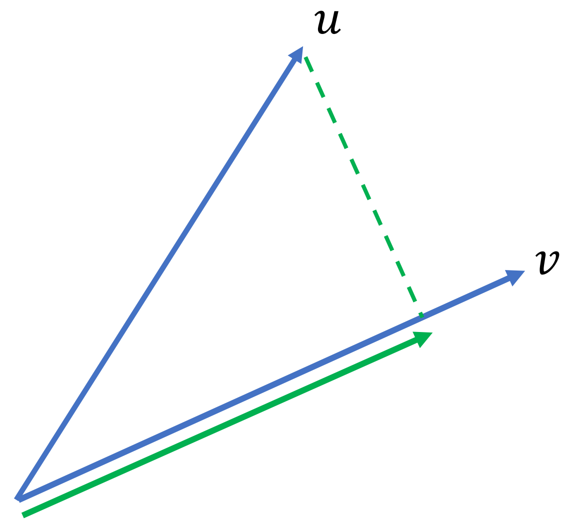
Or equivalently,

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$$

where $\|u\| = \sqrt{\langle u, u\rangle}$.

Proof 1 (case of $V = \mathbb{R}^d$):

- The left-hand-side is the length of v times the length of the projection of u onto v.
- Since orthogonal projection can't increase length, this is $\leq$ right-hand-side.



Important Properties

Proposition 3.1 (Cauchy-Schwartz): for any two vectors $u, v \in V$,

$$|\langle u, v\rangle|^2 \leq \langle u, u\rangle \cdot \langle v, v\rangle$$

Or equivalently,

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$$

where $\|u\| = \sqrt{\langle u, u\rangle}$.

And is a real number

Proof 2 (general case):

- If $v = 0_V$ then trivial, so can assume $v \neq 0_V$ and so $\langle v, v\rangle > 0$.
- Let $w = au + bv$. Get $0 \leq \langle w, w\rangle = |a|^2\langle u, u\rangle + |b|^2\langle v, v\rangle + \bar{a}b\langle u, v\rangle + \bar{b}a\langle v, u\rangle$.
- Set $a = \langle v, v\rangle, b = -\overline{\langle u, v\rangle}$. Get $\langle v, v\rangle^2\langle u, u\rangle + |\langle u, v\rangle|^2\langle v, v\rangle - 2|\langle u, v\rangle|^2\langle v, v\rangle$.
- $= \langle v, v\rangle(\langle v, v\rangle\langle u, u\rangle - |\langle u, v\rangle|^2)$. Since $\langle v, v\rangle > 0$, the 2nd term must be ≥ 0.

Important Properties

Exercise 3.2 *Prove that for any inner product space V and any $u, v, w \in V$*

$$\|u - w\| \leq \|u - v\| + \|v - w\| \ . \qquad \text{(triangle inequality)}$$

Proof:

- Let $x = u - v, y = v - w$. Want to show: $\|x + y\| \leq \|x\| + \|y\|$.
- Square both sides.
  - LHS = $\langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \leq \langle x, x\rangle + 2\|x\|\|y\| + \langle y, y\rangle = (\|x\| + \|y\|)^2$.

Important Properties

Exercise 3.2 *Prove that for any inner product space V and any $u, v, w \in V$*

$$\|u - w\| \;\leq\; \|u - v\| + \|v - w\| \; .$$

(triangle inequality)

This can be used to define convergence of sequences, and to define infinite sums and limits of sequences (which was not possible in an abstract vector space). However, it might still happen that the limit of a sequence of vectors in the vector space, which converges according to the norm defined by the inner product, may not converge to a vector in the space. Consider the following example.

Important Properties

Example 3.3 *Consider the vector space* $C([-1,1],\mathbb{R})$ *of continuous functions from* $[-1,1]$ *to* $\mathbb{R}$ *with the inner product defined by* $\langle f,g\rangle = \int_{-1}^{1} f(x)g(x)dx$. *Consider the sequence of functions:*

$$f_n(x) = \begin{cases} -1 & x \in \left[-1, \frac{-1}{n}\right) \\ nx & x \in \left[\frac{-1}{n}, \frac{1}{n}\right) \\ 1 & x \in \left[\frac{1}{n}, 1\right] \end{cases}$$

One can check that $\|f_n - f_m\|^2 = O(\frac{1}{n})$ for $m \geq n$. Thus, the sequence converges. However, the limit point is a discontinuous function not in the inner product space. To fix this problem, one can essentially include the limit points of all the sequences in the space (known as the completion of the space). An inner product space in which all (Cauchy) sequences converge to a point in the space is known as a Hilbert space. Many of the theorems we will prove will generalize to Hilbert spaces though we will only prove some of them for finite dimensional spaces.

Reminder: hwk1 due Wednesday