TTIC 31150/CMSC 31150
Mathematical Toolkit (Spring 2023)

Avrim Blum and Ali Vakilian

Lecture 5: The Real Spectral Theorem
Recap

• Eigenvectors and eigenvalues, eigenvectors of same eigenvalue form a subspace. Eigenvectors of different eigenvalues are linearly independent, inner products, norm, Cauchy-Schwartz.

• Gram-Schmidt orthogonalization, any finite-dimensional inner product space has an orthonormal basis.

• Properties of orthonormal bases: Fourier coefficients, Parseval’s identity

• Adjoint of a linear transform

• Reisz representation theorem. Use to prove that every linear transformation has a unique adjoint

• Self-adjoint linear operators: eigenvalues are real, and eigenvectors corresponding to distinct eigenvalues are orthogonal.
The Real Spectral Theorem

Theorem: every self-adjoint operator $\varphi : V \rightarrow V$ (which we know has real eigenvalues) has an orthonormal basis of eigenvectors (i.e., is “orthogonally diagonalizable”).

- E.g., square symmetric matrices over $\mathbb{R}^n$.

- Gives a nice way to view action of such operators. Say $\varphi$ has orthonormal eigenvectors $w_1, \ldots, w_n$ with associated eigenvalues $\lambda_1, \ldots, \lambda_n$. Then:

  For $v = \sum_i c_i w_i$, we have $\varphi(v) = \sum_i \lambda_i c_i w_i$.

  I.e., just stretching or shrinking in each “coordinate”.

Assume $V$ is finite-dimensional
The Real Spectral Theorem

Theorem: every self-adjoint operator $\varphi : V \to V$ (which we know has real eigenvalues) has an orthonormal basis of eigenvectors (i.e., is “orthogonally diagonalizable”).

Proof strategy:

1. Show that any such $\varphi$ has at least one eigenvalue.

2. Use (1) to prove the theorem.

We’ll do (2) first, then (1).
The Real Spectral Theorem

Theorem: every self-adjoint operator $\varphi: V \to V$ (which we know has real eigenvalues) has an orthonormal basis of eigenvectors (i.e., is “orthogonally diagonalizable”).

Proof part 2: induction on dimension of $V$.

- Base-case: $\dim(V) = 1$. By part (1), there is at least one eigenvalue and eigenvector, so just scale the eigenvector to be unit-length.

- Let $\dim(V) = k + 1$. Let $w$ be the eigenvector we are guaranteed by part (1) and let $W = \text{span}(\{w\})$. Let $W^\perp = \{v \in V: \langle v, w \rangle = 0\}$.

- Now, the idea to finish is to (a) show that $W^\perp$ is a subspace of $V$ of dimension $k$, (b) show that $\varphi$ restricted to $W^\perp$ is a self-adjoint operator on $W^\perp$ (and in particular maps $W^\perp$ to $W^\perp$), and (c) apply our inductive hypothesis to $W^\perp$ (which by design is orthogonal to $w$).
The Real Spectral Theorem

Theorem: every self-adjoint operator $\varphi: V \to V$ (which we know has real eigenvalues) has an orthonormal basis of eigenvectors (i.e., is “orthogonally diagonalizable”).

Assume $V$ is finite-dimensional.

Now, the idea to finish is to (a) show that $W^\perp$ is a subspace of $V$ of dimension $k$, (b) show that $\varphi$ restricted to $W^\perp$ is a self-adjoint operator on $W^\perp$ (and in particular maps $W^\perp$ to $W^\perp$), and (c) apply our inductive hypothesis to $W^\perp$ (which by design is orthogonal to $w$).

(a): If $\langle v_1, w \rangle = 0$ and $\langle v_2, w \rangle = 0$ then $\langle a_1v_1 + a_2v_2, w \rangle = 0$, so it’s a subspace.
Dimension is $k$ because a basis for $W^\perp \cup \{w\}$ is a basis for $V$. 
The Real Spectral Theorem

Theorem: every self-adjoint operator \( \varphi: V \to V \) (which we know has real eigenvalues) has an orthonormal basis of eigenvectors (i.e., is “orthogonally diagonalizable”).

Assume \( V \) is finite-dimensional.

\( \text{(b):} \) If \( \langle v, w \rangle = 0 \) want to show that \( \langle \varphi(v), w \rangle = 0 \).

- We can use the fact that \( \varphi \) is self-adjoint and \( w \) is an eigenvector.
- \( \langle \varphi(v), w \rangle = \langle v, \varphi(w) \rangle = \langle v, \lambda w \rangle = \lambda \langle v, w \rangle = 0 \).

Now, the idea to finish is to (a) show that \( W^\perp \) is a subspace of \( V \) of dimension \( k \), (b) show that \( \varphi \) restricted to \( W^\perp \) is a self-adjoint operator on \( W^\perp \) (and in particular maps \( W^\perp \) to \( W^\perp \)), and (c) apply our inductive hypothesis to \( W^\perp \) (which by design is orthogonal to \( w \)).
The Real Spectral Theorem

Assume $V$ is finite-dimensional

Theorem: every self-adjoint operator $\varphi: V \to V$ (which we know has real eigenvalues) has an orthonormal basis of eigenvectors (i.e., is “orthogonally diagonalizable”).

(c): Now, just apply induction.

- Let $\{w_1, ..., w_k\}$ be an orthonormal basis for $W^\perp$ of eigenvectors of $\varphi$ restricted to $W^\perp$.

- So, $\left\{w_1, ..., w_k, \frac{w}{\|w\|}\right\}$ is an orthonormal basis for $V$ of eigenvectors of $\varphi$.

• Now, the idea to finish is to (a) show that $W^\perp$ is a subspace of $V$ of dimension $k$, (b) show that $\varphi$ restricted to $W^\perp$ is a self-adjoint operator on $W^\perp$ (and in particular maps $W^\perp$ to $W^\perp$), and (c) apply our inductive hypothesis to $W^\perp$ (which by design is orthogonal to $w$).
The Real Spectral Theorem

Theorem: every self-adjoint operator \( \varphi: V \rightarrow V \) (which we know has real eigenvalues) has an orthonormal basis of eigenvectors (i.e., is “orthogonally diagonalizable”).

Proof strategy:

1. Show that any such \( \varphi \) has at least one eigenvalue.

2. Use (1) to prove the theorem.

Now, need to do (1).
Existence of eigenvalues

Let’s begin by assuming $V$ is over $\mathbb{C}$. Then won’t need self-adjointness.

Example: $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$
Existence of eigenvalues

Let’s begin by assuming $V$ is over $\mathbb{C}$. Then won’t need self-adjointness.

**Proposition 2.1** Let $V$ be a finite dimensional inner product space over $\mathbb{C}$ and let $\varphi : V \to V$ be a linear operator. Then $\varphi$ has at least one eigenvalue.

**Proof:** Let $\dim(V) = n$. Let $v \in V \setminus 0_V$ be any non-zero vector. Consider the set of $n + 1$ vectors $\{v, \varphi(v), \varphi^2(v), \ldots, \varphi^n(v)\}$ where $\varphi^i(v) = \varphi(\varphi^{i-1}(v))$. Since the dimension of $V$ is $n$, there must exist $c_0, \ldots, c_n \in \mathbb{C}$ not all 0 such that

$$c_0 \cdot v + c_1 \cdot \varphi(v) + \cdots + c_n \varphi^n(v) = 0_V.$$

For convenience, assume that $c_n \neq 0$, otherwise we can instead consider the sum to the largest $i$ such that $c_i \neq 0$. What we want to do now is to factor the expression above into a product of degree-1 terms. This is where working over $\mathbb{C}$ will be useful.
Existence of eigenvalues

Let’s begin by assuming $V$ is over $\mathbb{C}$. Then won’t need self-adjointness.

**Proposition 2.1** Let $V$ be a finite dimensional inner product space over $\mathbb{C}$ and let $\varphi : V \to V$ be a linear operator. Then $\varphi$ has at least one eigenvalue.

OK, so we have $c_0 v + c_1 \varphi(v) + \cdots + c_n \varphi^n(v) = 0_V$ with $c_n \neq 0$.

Let $P(x)$ denote the polynomial $c_0 + c_1 x + \cdots + c_n x^n$. Then the above can be written as $(P(\varphi))(v) = 0$, where $P(\varphi) : V \to V$ is a linear operator defined as

$$P(\varphi) := c_0 \cdot \text{id} + c_1 \cdot \varphi + \cdots + c_n \varphi^n,$$

with id used to denote the identity operator. Since $P$ is a degree-$n$ polynomial over $\mathbb{C}$, it can be factored into $n$ linear factors, and we can write $P(x) = c_n \prod_{i=1}^n (x - \lambda_i)$ for $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. This means that we can write

$$P(\varphi) = c_n (\varphi - \lambda_n \cdot \text{id}) \cdots (\varphi - \lambda_1 \cdot \text{id}).$$
Existence of eigenvalues

Let’s begin by assuming $V$ is over $\mathbb{C}$. Then won’t need self-adjointness.

**Proposition 2.1** Let $V$ be a finite dimensional inner product space over $\mathbb{C}$ and let $\varphi : V \to V$ be a linear operator. Then $\varphi$ has at least one eigenvalue.

OK, so we have $P(\varphi) = c_n (\varphi - \lambda_n \cdot id) \ldots (\varphi - \lambda_1 \cdot id)$, and $P(\varphi)(v) = 0$.

Let $w_0 = v$ and define $w_i = \varphi(w_{i-1}) - \lambda_i \cdot w_{i-1}$ for $i \in [n]$. That is, we are working through the computation of $P(\varphi)(v)$ from right to left. Note that $w_0 = v \neq 0_V$ and $w_n = P(\varphi)(v) = 0_V$. Let $i^*$ denote the largest index $i$ such that $w_i \neq 0_V$. Then, we have

$$0_V = w_{i^*+1} = \varphi(w_{i^*}) - \lambda_{i^*+1} \cdot w_{i^*}.$$ 

This means that $w_{i^*}$ is an eigenvector of $\varphi$ with eigenvalue $\lambda_{i^*+1}$. 
Existence of eigenvalues

Now, what about when $V$ is over $\mathbb{R}$?

- Can do the same argument, except $P$ now factors into linear and quadratic terms.
- Just need to show that we hit 0 in one of the linear terms, and not one of the irreducible quadratic terms.
- Specifically, want to show we don’t get an equation of the form:
\[
0_V = \varphi^2(w_i^*) + b\varphi(w_i^*) + cw_i^*, \text{ with } b^2 < 4c
\]

This is where self-adjointness comes in.
Existence of eigenvalues

Now, what about when $V$ is over $\mathbb{R}$?

- Want to show we don’t get an equation of the form:
  \[ 0_V = \varphi^2(w_i^*) + b\varphi(w_i^*) + cw_i^*, \text{with } b^2 < 4c \]

\[
\langle w_i^*, \varphi^2(w_i^*) + b\varphi(w_i^*) + cw_i^* \rangle = \langle w_i^*, \varphi^2(w_i^*) \rangle + b\langle w_i^*, \varphi(w_i^*) \rangle + c\langle w_i^*, w_i^* \rangle \\
= \langle \varphi(w_i^*), \varphi(w_i^*) \rangle + b\langle w_i^*, \varphi(w_i^*) \rangle + c\langle w_i^*, w_i^* \rangle \\
= \| \varphi(w_i^*) \|^2 + b\langle w_i^*, \varphi(w_i^*) \rangle + c\| w_i^* \|^2 \\
\geq \| \varphi(w_i^*) \|^2 - |b|\| w_i^* \| \| \varphi(w_i^*) \| + c\| w_i^* \|^2 \\
= \left( \| \varphi(w_i^*) \| - \frac{|b|\| w_i^* \|}{2} \right)^2 + \left( c - \frac{b^2}{4} \right) \| w_i^* \|^2 \\
> 0.
\]

So, the quadratic term can’t be 0.
**Raleigh Quotients**

**Definition 3.1** Let $\phi : V \to V$ be a self-adjoint linear operator and $v \in V \setminus \{0_V\}$. The Rayleigh quotient of $\phi$ at $v$ is defined as

$$\mathcal{R}_\phi(v) := \frac{\langle v, \phi(v) \rangle}{\|v\|^2}.$$  

We can equivalently write $\mathcal{R}_\phi(v) = \langle \hat{v}, \phi(\hat{v}) \rangle$ for $\hat{v} = v/\|v\|$.

In other words, it is the length of the projection of $\phi(\hat{v})$ onto $\hat{v}$.

If $v$ was an eigenvector, then this would be the eigenvalue.
Raleigh Quotients

**Definition 3.1** Let $\varphi : V \to V$ be a self-adjoint linear operator and $v \in V \setminus \{0_V\}$. The Rayleigh quotient of $\varphi$ at $v$ is defined as

$$R_\varphi(v) := \frac{\langle v, \varphi(v) \rangle}{\|v\|^2}.$$  

We can equivalently write $R_\varphi(v) = \langle \hat{v}, \varphi(\hat{v}) \rangle$ for $\hat{v} = v/\|v\|$.

**Proposition 3.2** Let $\dim(V) = n$ and let $\varphi : V \to V$ be a self-adjoint operator with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then,

$$\lambda_1 = \max_{v \in V \setminus \{0_V\}} R_\varphi(v) \quad \text{and} \quad \lambda_n = \min_{v \in V \setminus \{0_V\}} R_\varphi(v)$$

So, the vector $v$ such that applying $\varphi$ gives the largest “stretch” in the $\hat{v}$ direction is the eigenvector of largest eigenvalue, and likewise for the eigenvector of smallest eigenvalue.
Raleigh Quotients

**Definition 3.1** Let \( \varphi : V \to V \) be a self-adjoint linear operator and \( v \in V \setminus \{0_V\} \). The Rayleigh quotient of \( \varphi \) at \( v \) is defined as

\[
R_{\varphi}(v) := \frac{\langle v, \varphi(v) \rangle}{\|v\|^2}.
\]

We can equivalently write \( R_{\varphi}(v) = \langle \hat{v}, \varphi(\hat{v}) \rangle \) for \( \hat{v} = v / \|v\| \).

**Proposition 3.2** Let \( \text{dim}(V) = n \) and let \( \varphi : V \to V \) be a self-adjoint operator with eigenvalues \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \). Then,

\[
\lambda_1 = \max_{v \in V \setminus \{0_V\}} R_{\varphi}(v) \quad \text{and} \quad \lambda_n = \min_{v \in V \setminus \{0_V\}} R_{\varphi}(v).
\]

Proof: Let \( w_1, \ldots, w_n \) be an orthonormal basis of eigenvectors with eigenvalues \( \lambda_1, \ldots, \lambda_n \). Let \( \hat{v} = \sum_i c_i w_i \). Then \( \langle \hat{v}, \varphi(\hat{v}) \rangle = \langle \sum_i c_i w_i, \sum_i \lambda_i c_i w_i \rangle = \sum_i \lambda_i |c_i|^2 \). Since \( \sum_i |c_i|^2 = 1 \), this is a weighted average of the \( \lambda_i \)'s, and so is maximized at \( c_1 = 1 \), and minimized at \( c_n = 1 \).
Raleigh Quotients

**Definition 3.1** Let $\varphi : V \to V$ be a self-adjoint linear operator and $v \in V \setminus \{0_V\}$. The Rayleigh quotient of $\varphi$ at $v$ is defined as

$$R_\varphi(v) := \frac{\langle v, \varphi(v) \rangle}{\|v\|^2}.$$

We can equivalently write $R_\varphi(v) = \langle \hat{v}, \varphi(\hat{v}) \rangle$ for $\hat{v} = v / \|v\|$. 

**Extension / Generalization:**

**Proposition 3.3 (Courant-Fischer theorem)** Let $\dim(V) = n$ and let $\varphi : V \to V$ be a self-adjoint operator with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then,

$$\lambda_k = \max_{\dim(S) = k} \min_{v \in S \setminus \{0_V\}} R_\varphi(v) = \min_{\dim(S) = n - k + 1} \max_{v \in S \setminus \{0_V\}} R_\varphi(v).$$
Positive Semidefiniteness

**Definition 3.4** Let \( \varphi : V \rightarrow V \) be a self-adjoint operator. \( \varphi \) is said to be positive semidefinite if \( R_\varphi(v) \geq 0 \) for all \( v \neq 0 \). \( \varphi \) is said to be positive definite if \( R_\varphi(v) > 0 \) for all \( v \neq 0 \).

**Proposition 3.5** Let \( \varphi : V \rightarrow V \) be a self-adjoint linear operator. Then the following are equivalent:

1. \( R_\varphi(v) \geq 0 \) for all \( v \neq 0 \).

2. All eigenvalues of \( \varphi \) are non-negative.

3. There exists \( \alpha : V \rightarrow V \) such that \( \varphi = \alpha^* \alpha \).

Part of argument: if \( \varphi = \alpha^* \alpha \) then \( \langle v, \varphi(v) \rangle = \langle v, \alpha^*(\alpha(v)) \rangle = \langle \alpha(v), \alpha(v) \rangle \geq 0 \). This also means that if \( v \) is an eigenvector, its eigenvalue must be non-negative.

The decomposition of a positive semidefinite operator in the form \( \varphi = \alpha^* \alpha \) is known as the Cholesky decomposition of the operator. Note that if we can write \( \varphi \) as \( \alpha^* \alpha \) for any \( \alpha : V \rightarrow W \), then this in fact also shows that \( \varphi \) is self-adjoint and positive semidefinite.