**CSE 840: Computational Foundations of Artificial Intelligence** September 25, 2023

# Symmetric Matrices to Characterization of Eigenvalues

*Instructor: Vishnu Boddeti* *Scribe: Joe Romain, Aidan McCoy, Tashfain Ahmed*

## 1 Symmetric Matrices

### Definition

Matrix A $\in \mathbb{R}^{n \times n}$ is called a symmetric matrix if $A = A^T$
Matrix $A \in \mathbb{C}^{n \times n}$ is called a Hermitian matrix if $A = \bar{A}^T$

### Proposition

Let A$\in \mathbb{C}^{n \times n}$ be Hermetian then all eigen values of A are real valued.
Eigenvectors that corresponds to distinct eigenvalue are orthogonal.

#### Proof

$\lambda\langle x, x\rangle = \langle \lambda x, x\rangle = \langle Ax, x\rangle$
Because Ax $= \lambda x$
$\langle x, Ax\rangle = \langle x, \lambda x\rangle = \bar{\lambda}\langle x, x\rangle$

$\lambda = \bar{\lambda} \in \mathbb{R}$ (Unless x=0 vector, it has to be real)

Consider: $(\lambda_1, x_1)$ and $(\lambda_2, x_2)$ are eigenvalue-eigenvector pairs of $A$

$\lambda_1\langle x_1, x_2\rangle = \langle \lambda_1 x_1, x_2\rangle = \langle Ax_1, x_2\rangle$
$= \langle x_1, Ax_2\rangle = \langle x_1, \lambda_2 x_2\rangle$

$= \bar{\lambda_2}\langle x_1, x_2\rangle$

$0 = \lambda_1\langle x_1, x_2\rangle - \lambda_2\langle x_1, x_2\rangle$
$0 = (\lambda_1 - \lambda_2)\langle x_1, x_2\rangle$

$$\text{Either } \lambda_1 = \lambda_2$$
$$\text{or if } \lambda_1 \neq \lambda_2, \text{ then } \langle x_1, x_2\rangle = 0$$
$$\Rightarrow x_1 \perp x_2$$

## Definition

An operator T in L(V) on a Pre-Hilbert space V is called self-adjoint if
T $\langle v, w \rangle = \langle v, Tw \rangle$

Sometimes it's called a Hermitian Operator (on $\mathrm{C}^n$) or a Symmetric Operator (on $\mathbb{R}^n$).

### Remarks:

Over $\mathrm{C}^n$, self adjoint operators are represented by Hermetian Matrices. Similarly.
Over$\mathbb{R}^n$, self adjoint operators are represented by Symmetric Matrices.

## Proposition

Let $T$ in $\mathcal{L}(V)$ be self adjoint. Then, T has atleast one eigen value and it's real value (holds both on $\mathrm{C}^n$and$\mathbb{R}^n$)

### Proof(Sketch)

Let $n := \dim V$ and choose $v \neq 0$.

Consider the vectors $v, Tv, T^2v, \ldots, T^nv$. These vectors are linearly dependent ($n+1$ vectors, $\dim V = n$). There exist coefficients $a_0, a_1, \ldots, a_n$.

So, vectors with coefficients are:

$$a_0 v + a_1 Tv + \ldots + a_n T^n v$$

Consider the polynomial with these coefficients:

$$P(x) = a_0 + a_1 x + \ldots + a_n x^n = 0 = \underbrace{C(x^2 + b_1 x + c_1)\ldots(x^n + b_m x + c_m)}_{\text{(quadratic terms)}} \times \underbrace{(x - \lambda_1)\ldots(x - \lambda_m)}_{\text{(linear terms)}}$$

Replace $x$ by $T$:

$$0 = a_0 + a_1 Tv + \ldots + a_n T^n v = \underbrace{C(T^2 + b_1 T + c_1)\ldots(T^n + b_m T + c_m)}_{\text{(quadratic terms)}} \times \underbrace{(T - \lambda_1 I)\ldots(T - \lambda_m I)}_{\text{(linear terms)}} v$$

Now, we can show that the quadratic terms are invertible, and we are left with at least one linear factor:

$$0 = (T - \lambda_1 I)\ldots(T - \lambda_n I)v$$

There needs to exist at least one $i$ such that $(T - \lambda_i I)$ is not invertible.
So $(T - \lambda_i I)v = 0$.
This implies $Tv = \lambda_i v$.
which means $\lambda_i$ is an eigenvalue of $T$.

## 2 Spectral Theorem for Symmetric/ Hermitian Matrices

**Theorem 1** *A symmetric matrix $A \in \mathbb{R}^{n\times n}$ is orthognally diagonalizable: there exists an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ and a diagonal matrix $D \in \mathbb{R}^{n\times n}$ s.t.*

$$A = QDQ^\top, D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$
$$= \sum_{i=1}^{n} \lambda_i \underbrace{q_i q_i^\top}_{\hookrightarrow \; rank-1\, matrices}$$

**Theorem 2** *A Hermitian matrix $A \in \mathbb{C}^{n\times n}$ is unitarily diagonalizable: there exists a unitary matrix $U$ and a diagonal matrix $D$. s.t.*

$$A = UD\bar{U}^\top$$

*The entries of $D$ are real-valued.*

## 3 Positive Definite Matrices

**Definition 3** *A matrix $A \in \mathbb{R}^{n\times n}$ is called a* ***positive definite*** *(PD) if $\forall x \in \mathbb{R}^n$,*

$$x \neq 0 \quad x^\top Ax > 0$$

*For* ***positive semi-definite*** *(PSD) $\forall x \in \mathbb{R}^n$ $x \neq 0, \quad x^\top Ax \geq 0$*

**Definition 4** *A matrix $A \in \mathbb{C}^{n\times n}$ is called a Gram matrix if there exists a set of vectors $v_1, \dots v_n \in \mathbb{C}^n$ such that $a_{ij} = \langle v_i, v_j \rangle$. Note: Gram matrices are Hermitian (Similarly on $\mathbb{R}^{nan}$ , then Gram matrices are symmetric).*

$$G = V^\top V, \quad V = \begin{bmatrix} 1 & & 1 \\ V_1 & \ddots & V_n \\ 1 & & 1 \end{bmatrix}$$
$$CV^\top$$

⚠ Over $\mathbb{C}$, we have that $PD \Rightarrow$ self adjoint.

Over $\mathbb{R}$, this is **not** true!

$\Rightarrow$ there are matrices which are $PD$ but not symmetric.

Example:

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$x^\top Ax = x_1^2 + x_2^2 > 0$$

$\rightarrow$ So $A$ is $PD$ but not symmetric

$\rightarrow$ over $\mathbb{C}$, the same matrix is not $PD$ since $x_1^2 + x_2^2$ can be negative!

**Theorem 5** *$A \in \mathbb{C}^{n\times n}$ Hermitian. Then equivalent:*

*(i) $A$ is $PSD$ ($PD$)*

*(ii) All eigenvalues of $A$ are $\geq 0 \quad (> 0)$*

*(iii) The mapping $\langle \cdot, \cdot \rangle_A : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ with*

$$\langle x, y \rangle_A := \bar{y}^\top Ax$$

*satisfies all properties of an inner product except one: if $\langle x, x \rangle_A = 0$, this does not imply $x = 0$*

*(This mapping is an inner product)*

*(iv) $A$ is a Gram matrix of $n$ vectors which are not necessarily linearly independent (which are linearly independent).*

$$a_{ij} = \langle x_i, x_j \rangle$$

# 4 Roots of PSD matrices

**Theorem 6** *Let $A \in \mathbb{R}^{n\times n}$ be symmetric, PSD. Then there exists a matrix $B \in \mathbb{R}^{n\times n}, B$ is PSD such that $A = B^2$. Sometimes $B$ is called the square root of $A$,*

$$B = (A)^{1/2}$$

**Proof:** Spectral theorem $\Rightarrow$

$$A = UDU^\top, D \text{ diagonal}$$

$PSD \Rightarrow$ eigenvalues $\Rightarrow 0$

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}, \lambda_i \geq 0$$

Define

$$\sqrt{D} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{\lambda_n} \end{bmatrix}$$

and set

$$B := U\sqrt{D}U^\top$$

This matrix satisfies the property that $B^2 = A$

## 5 Variational Characterization of Eigenvalues

**Definition 7** *Let $A \in \mathbb{R}^{nxn}$ be a symmetric matrix.*

$$R_A : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$$

$$x \mapsto \frac{x^T Ax}{x^T x}$$

*is called the Rayleigh coefficient.*

**Proposition 8** *Let A be symmetric, let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ be the eigenvalues and $v_1, v_2, ..., v_n$ the eigenvectors of A. Then:*

$$\min_{x \in \mathbb{R}^n} R_A(x) = \min_{||x||=1} x^T Ax = \lambda_1, \text{ attained at } x = v_1$$

$$\max_{x \in \mathbb{R}^n} R_A(x) = \max_{||x||=1} x^T Ax = \lambda_n, \text{ attained at } x = v_n$$

**Intuition:**

Assume A is expressed in terms of the basis $v_1, \ldots, v_n$

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}$$

Let y be a vector, also represented in the same basis.

y $= y_1 v_1 + y_2 v_2 + \cdots + y_n v_n$

$y^T Ay = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$

Among the vectors $\begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \\ \dots \end{pmatrix} \dots \begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \end{pmatrix}$, the smallest result of $y^T A y$ would be given by the vector $\begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}$, $(v_1)$, and the value would be $\lambda_1$.

**More general proof (sketch):**

Assume we start with the standard basis. Let Q= $\begin{pmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_n \\ | & | & & | \end{pmatrix}$ be the basis transformation.

Observe: Q is orthogonal, we have

$A = Q^T \Lambda Q$, where $\Lambda$ is diagonal.

For a vector x= $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ in the original basis, we now consider $y := Q^T x$.

$R_A(y) = \frac{y^T A y}{y^T y}$

$= \frac{(Q^T x)^T A (Q^T x)}{(Q^T x)^T (Q^T x)}$

$= \frac{x^T Q Q^T \Lambda Q Q^T x}{x^T Q Q^T x}$ $Note : (Q^T x)^T = x^T Q$

$= \frac{x^T \Lambda x}{x^T x}$ Note: Because Q is orthogonal, $QQ^T = I$

$= \frac{\lambda_1 x_1^2 + \dots + \lambda_n x_n^2}{||x||}$

$$\min_{||y||=1} R_A(y) = \min_{||x||=1} \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$$

Note: Q is orthogonal, which means norms are preserved

This minimum is attained for x= $\begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}$, that is $y = Q^T x = v_1$, with value

$$\min_{||y||=1} R_A(y) = \lambda_1$$

**Proposition 9** *Consider the problem*

$$\min_{\substack{||x||=1 \\ x \perp v_1}} R(x).$$

*The solution to this problem is* $x = v_2, R(x) = \lambda_2$.

**Intuition:** Consider operator A restricted to the space $V_1^{\perp} := (span\{v_1\})^{\perp}$. We know that on this space, A is invariant and symmetric, so we can apply Rayleigh to this "smaller" space.

$$V_1^{\perp} = span\{v_2, v_3, \dots, v_n\}$$

If we apply Rayleigh to $V_1^{\perp}$, we get the solution $\lambda_2, v_2$.

**Theorem 10** *(Min-Max Theorem):*

*Let* $A \in \mathbb{R}^{nxn}$ *be symmetric, with eigenvalues* $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. *Then*

$$\lambda_k = \min_{\substack{U subspace \\ dim U = k}} \max_{x \in U \setminus \{0\}} R_A(x)$$

$$= \max_{\substack{U subspace \\ dim U = n-k+1}} \min_{x \in U \setminus \{0\}} R_A(x)$$

**Intuition:** for k=3,

Consider the subspace U, spanned by $v_1, v_2, v_3$. As we saw before,

$$\max_{x \in U} R_A(x) = \lambda_3, \ \textit{attained by } v_3.$$

Consider another subspace U, spanned by $v_9, v_{10}, v_{11}$.

$$\max_{x \in U} R_A(x) = \lambda_{11}$$