Symmetric Matrices

Def: A matrix $A \in \mathbb{R}^{n \times n}$ is called symmetric if $A = A^T$

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

Proposition: Let $A \in \mathbb{C}^{n \times n}$ be Hermitian. Then all eigenvalues of A are real-valued. Eigenvectors that correspond to distinct eigenvalues are orthogonal.

Proof: λ eigenvalue of A with eigenvector x. Then.

$Ax = \lambda x$

Hermitian

$$\underline{\lambda \langle x, x \rangle} = \langle \lambda x, x \rangle = \langle Ax, x \rangle =$$

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

$$\Rightarrow \underline{\lambda = \bar{\lambda}} \in \mathbb{R} \quad (\text{unless } x = 0 \text{ vector})$$

$\hookrightarrow$ λ has to be real

$(\lambda_1, x_1), (\lambda_2, x_2)$ are eigenvalue - eigenvector pairs of A. Then

$$\lambda_1 \langle x_1, x_2 \rangle = \langle \lambda_1 x_1, x_2 \rangle = \langle A x_1, x_2 \rangle = \langle x_1, A x_2 \rangle$$

$$= \langle x_1, \lambda_2 x_2 \rangle = \bar{\lambda}_2 \langle x_1, x_2 \rangle$$

$$\hookrightarrow \bar{\lambda}_2 = \lambda_2$$

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

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

$\Rightarrow$ either $\lambda_1 = \lambda_2$

or if $\lambda_1 \neq \lambda_2$, then $\langle x_1, x_2 \rangle = 0$

$$\Rightarrow x_1 \perp x_2$$

$\blacksquare$

Def: An operator $T \in \mathcal{L}(V)$ on a pre-Hilbert space V is called <u>self-adjoint</u> if

$$\langle Tv, w \rangle = \langle v, Tw \rangle.$$

Sometimes it is called a Hermitian operator (on $\mathbb{C}^n$)

Symmetric operator (on $\mathbb{R}^n$)

Remark: Over $\mathbb{C}^n$, self-adjoint operators are represented by Hermitian matrices. On $\mathbb{R}^n$, self-adjoint operator is represented by a symmetric matrix.

Proposition: $T \in \mathcal{L}(V)$ self-adjoint. Then T has at least one eigenvalue, and it is real-valued. (holds both on $\mathbb{C}^n$ and $\mathbb{R}^n$).

Proof (sketch): $n := \dim V$. Chose $v \neq 0$, and consider,

$$v, Tv, T^2 v, \dots T^n v$$

These vectors have to be linearly dependent ($n+1$ vectors, $\dim = n$).

There exists $a_0, a_1 \dots a_n$:

$$a_0 v + a_1 T v + \dots + a_n T^n v = 0.$$

Consider polynomial with these coefficients:

$$a_0 + a_1 x + a_2 x^2 + \dots + a_n x_n = 0$$

$$= c\underbrace{(x^2 + b_1 x + c_1) \dots (x^2 + b_m x + c_m)}_{\text{quadratic terms}}\underbrace{(x - \lambda_1) \dots (x - \lambda_m)}_{\text{linear terms}}$$

Replace the x by T:

$$0 = (a_0 + a_1 Tv + \dots a_n T^n v) = (c \underbrace{(\dots\dots)}_{\text{quadratic}} \underbrace{(\dots)}_{\text{linear}})v$$

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

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

There needs to exist at least one 'i' such that $(T - \lambda_i I)$ is not invertible

So. $(T - \lambda_i I) v = 0 \implies Tv = \lambda_i v$

$\implies \lambda_i$ is a eigenvalue of T.

$\blacksquare$

Spectral Theorem for Symmetric / Hermitian Matrices

Theorem: 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.

$$A = Q D Q^T \qquad D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

$$= \sum_{i=1}^{n} \lambda_i \underbrace{q_i q_i^T}_{\text{rank-1 matrices}}$$

Theorem: 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.

$$A = U D \bar{U}^T$$

the entries of D are real-valued.

Positive Definite Matrices

Def: A matrix $A \in \mathbb{R}^{n\times n}$ is called a positive definite (PD) if $\forall x \in \mathbb{R}^n$, $x \neq 0$ $\quad x^T A x > 0$

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

Def: 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 $\mathbb{R}^{n\times n}$, then Gram matrices are symmetric).

$$G = V^T V, \quad V = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}$$

$$C = V V^T$$

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

Over $\mathbb{R}$, this is <u>not</u> true!

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

<u>Example:</u> $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$

$$x^T A x = x_1^2 + x_2^2 > 0$$

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

(i) A is PSD (PD)

(ii) All eigvenvalues of A are ≥ 0 (>0)

(iii) the mapping $\langle \cdot, \cdot \rangle_A : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}$ with

$$\langle x, y \rangle_A := \bar{y}^T A x$$

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).

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

Roots of PSD matrices

Theorem: 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 $\Rightarrow$

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

- PSD $\Rightarrow$ eigenvalues ≥ 0

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix}, \quad \lambda_i \geq 0$$

Define $\sqrt{D} = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix}$ and set

$$B := U\sqrt{D}\,U^T.$$

$\blacksquare$

Variational Characterization of Eigenvalues

Def: Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix.

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

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

is called the Rayleigh coefficient.

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

Then:

$$\min_{x \in \mathbb{R}^n} R_A(x) = \min_{\|x\|=1} x^T A x$$

$$= \lambda_1 , \text{ attained at } x = v_1$$

$$\max_{x \in \mathbb{R}^n} R_A(x) = \max_{\|x\|=1} x^T A x$$

$$= \lambda_n , \text{ attained at } x = v_n$$

<u>Intuition</u>: Assume A is expressed in terms of the basis $V_1 \dots V_n$

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

Let Y be a vector, also represted in the same basis.

$$Y = y_1 v_1 + y_2 v_2 + \dots + y_n v_n$$

$$Y^T A Y = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$$

Among the vectors $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} \dots \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$

the smallest result of $Y^T A Y$ would be given by the vector $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \to V_1$, and the value would be λ_1.

More general proof (sketch):

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

Observe: Q is orthogonal, we have

$$A = Q^T \Lambda Q \quad \text{with } \Lambda \text{ diagonal}$$

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

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

$(Q^T x)^T = x^T Q$

$$= \frac{x^T \underbrace{\cancel{Q Q^T}}_{I} \Lambda \underbrace{\cancel{Q Q^T}}_{I} x}{x^T \underbrace{\cancel{Q Q^T}}_{I} x} = \frac{x^T \Lambda x}{x^T x} = \frac{\lambda_1 x_1^2 + \ldots + \lambda_n x_n^2}{\|x\|}$$

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

Q is orthogonal matrix, which preserves norms.

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

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

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

$$V_1^\perp = \text{span}\{V_2, V_3 \ldots V_n\}$$

If we apply Rayleigh to $V_1^\perp$, then we get the solution λ_2, v_2.

Theorem: (Min-Max-Theorem)

$A \in \mathbb{R}^{n \times n}$ symmetric, eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Then

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

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

Intuition: for $k = 3$

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

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

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

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