2

Diagonalization

If an $n \times n$ matrix A happens to have n linearly independent eigenvectors, then it can be written (or "diagonalized") as

$$A = T \Lambda T^{-1}$$

where

- T is an $n \times n$ invertible matrix, and
- Λ is an $n \times n$ diagonal matrix.

Construction:

- Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A. (We do not necessarily assume they are distinct.)
- Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ the the corresponding eigenvectors. (We do assume that these are linearly independent.)
- For all $i = 1, 2, \ldots, n$, we know that $A\vec{v}_i = \lambda_i \vec{v}_i$

Diagonalization - 2

$A\vec{v}_i = \lambda_i \vec{v}_i, \quad i = 1, 2, \cdots, n.$

- We can stack these n equations in the form of a matrix equation:

$$A \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \cdots & \lambda_n\vec{v}_n \\ | & | & & | \end{bmatrix}$$

that is,

$$A \underbrace{\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}}_{T} = \underbrace{\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}}_{\Lambda}$$

$AT = T\Lambda$

- Because the $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ are linearly independent, then T must be invertible. Thus,

$\underbrace{ATT^{-1}}_{=A} = T\Lambda T^{-1}$

$$A = T\Lambda T^{-1}.$$

Diagonalization of Hermitian matrices

Lemma (Eigenvalues of a Hermitian matrix)

If $A = A^H$, then all eigenvalues of A are real-valued *(This is true even if A itself has complex entries.)*

Proof:

- Let λ be an eigenvalue of A and let $\vec{x}$ be an eigenvector corresponding to λ.
- Then

$$\langle \underbrace{A\vec{x}}_{\lambda\vec{x}}, \vec{x}\rangle = \langle \lambda\vec{x}, \vec{x}\rangle = \lambda\langle \vec{x}, \vec{x}\rangle$$

- But also,

$$\langle \vec{x}, A\vec{x}\rangle = \langle \vec{x}, \lambda\vec{x}\rangle = \lambda^*\langle \vec{x}, \vec{x}\rangle$$

$= \quad \Rightarrow \lambda = \lambda^*$

$\Rightarrow \lambda$ is real.

- Since $A = A^H$, then $\langle A\vec{x}, \vec{x}\rangle$ must equal $\langle \vec{x}, A\vec{x}\rangle$, and so this implies that λ must equal λ^*.

  $= \langle \vec{x}, A^H\vec{x}\rangle$ //

- Therefore, λ must be real. $\square$

This lemma does not mean that all the eigenvalues must be distinct (only that they must be real). So what we can say about the eigenvectors? Will they be linearly independent?

Diagonalization of Hermitian matrices - 2

Lemma (Eigenvectors of a Hermitian matrix)

If $A = A^H$, then there exist a set of n orthonormal *eigenvectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ such that*

$$A\vec{v}_i = \lambda_i \vec{v}_i$$

for all $i = 1, 2, \ldots, n$.

This result holds even if there are repeated eigenvalues, but it uses the assumption that $A = A^H$.

Diagonalization of Hermitian matrices - 3

Diagonalization of a Hermitian matrix:

If T is a unitary matrix,

$T^H T = T T^H = I$.

- Suppose $\overset{n\times n}{A} = A^H$.
- Then choosing an orthonormal set of eigenvectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ and letting $T = [\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n]$ as before, we have

$$A = T\Lambda T^{-1}$$

- However, since the $\{\vec{v}_i\}$ are orthonormal, then T is unitary. Therefore, $T^{-1} = T^H$ and so

$$A = T\Lambda T^H$$

Note: if A is real, it is possible to choose T real and have $A = T\Lambda T^T$. transpose.

Diagonalization of Hermitian matrices - 4

$\det(A-\lambda I)$
$= \det\left(\begin{bmatrix} 3/2-\lambda & 1/2 \\ 1/2 & 3/2-\lambda \end{bmatrix}\right)$

Example:

- Let

$$A = A^H = \begin{bmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{bmatrix}$$

$= (3/2-\lambda)^2 - \frac{1}{4} = 0.$

$\Rightarrow 3/2-\lambda = \pm\frac{1}{2}$

  Then $\lambda_1 = 2$ and $\lambda_2 = 1$, both of which are real since $A = A^H$. $\Downarrow$

$\lambda_1 = 2,\ \lambda_2 = 1.$

- We can derive

$$\vec{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

- Thus, $A = T\Lambda T^H$, where

$T = [\vec{v}_1\ \vec{v}_2] = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},\ \Lambda = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$

$\begin{cases} \langle \vec{v}_1, \vec{v}_2 \rangle = 0. \\ \|\vec{v}_1\|_2 = \|\vec{v}_2\|_2 = 1. \end{cases}$

Positive definite matrices: Definition

Let A be a square (say, $n \times n$), Hermitian symmetric matrix.

$\rightarrow A \succ B \Leftrightarrow A - B \succ 0$

Recall that we say that A is *positive definite* if PD

$$\vec{x}^H A \vec{x} > 0 \qquad A \succ 0$$

" $\succ$ " is different from " $>$ ".

holds for all non-zero $\vec{x} \in \mathbb{R}^n$ (or $\mathbb{C}^n$).

$a > b$

$3 > 1$

Similarly, we say that A is *positive semi-definite* if PSD

$$\vec{x}^H A \vec{x} \geq 0 \qquad A \succeq 0$$

holds for all non-zero $\vec{x} \in \mathbb{R}^n$ (or $\mathbb{C}^n$).

Such matrices are also called symmetric, positive (semi-)definite.

$A = 3$ (1×1) PD

$A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \leftarrow$ PD

Positive definite matrices: Eigenvalues

If $A = A^H$, we already know the eigenvalues of A are real.

Furthermore, $\boxed{\text{if } A \text{ is positive definite, then all eigenvalues of } A \text{ are positive.}}$

- Proof: Let $\vec{v}$ be an eigenvector of A and let λ be the corresponding eigenvalue. Assume $\vec{v} \neq \vec{0}$. Then, because A is positive definite,

$$\vec{v}^H \underbrace{A\vec{v}}_{=\lambda\vec{v}} > 0$$

Substituting,

$$\vec{v}^H(\lambda\vec{v}) > 0 \Rightarrow \lambda \underbrace{\vec{v}^H\vec{v}}_{=\|\vec{v}\|_2^2>0} > 0 \Rightarrow \lambda > 0.$$

because $\|\vec{v}\| > 0$.

Similarly, $\boxed{\text{if } A \text{ is positive semi-definite, then all eigenvalues of } A \text{ are nonnegative.}}$

Positive definite matrices: Weighted inner products

Positive definite matrices can be used to define variations on the standard ℓ_2 inner product.

In particular, suppose A is a symmetric, positive definite matrix. Then

Hermitian

$$\langle \vec{x}, \vec{y} \rangle_A := \vec{y}^H A \vec{x}$$

$\langle \vec{x}, \vec{y} \rangle = \vec{y}^H \vec{x}$

$\Downarrow$

$A = I$

$\langle \vec{x}, \vec{y} \rangle_I = \vec{y}^H \vec{x}.$

defines a valid inner product on $\mathbb{C}^n$.

Consequently,

$$\|\vec{x}\|_A = \sqrt{\langle \vec{x}, \vec{x} \rangle_A} = \sqrt{\vec{x}^H A \vec{x}}$$

defines a valid induced norm on $\mathbb{C}^n$.

Positive definite matrices: Geometry

Suppose A is symmetric Hermitian and positive definite.

Consider the optimization problems

$$\max_{\vec{x}\in\mathbb{C}^n} \frac{\|\vec{x}\|_A^2}{\|\vec{x}\|_2^2} = \boxed{\max_{\substack{\vec{x}\in\mathbb{C}^n \\ \|\vec{x}\|_2=1}} \vec{x}^H A\vec{x}} \quad \text{and} \quad \min_{\vec{x}\in\mathbb{C}^n} \frac{\|\vec{x}\|_A^2}{\|\vec{x}\|_2^2} = \min_{\substack{\vec{x}\in\mathbb{C}^n \\ \|\vec{x}\|_2=1}} \vec{x}^H A\vec{x}.$$

The maximum value of the first problem is given by $\lambda_{max}(A)$ and occurs when $\vec{x}$ equals the corresponding eigenvector of A.

Similarly, the minimum value of the second problem is given by $\lambda_{min}(A)$ and occurs when $\vec{x}$ equals the corresponding eigenvector of A.

Positive definite matrices: Geometry - 2

$\langle \vec{x}, \vec{v}_i \rangle = \langle \sum_{j=1}^{n} \alpha_j \vec{v}_j , \vec{v}_i \rangle$

$= \sum_{j=1}^{n} \alpha_j \underbrace{\langle \vec{v}_j , \vec{v}_i \rangle}$

$= 1, j = i$

$= 0, j \neq i$

$= \alpha_i$

Proof (for the maximization problem):

- Since $A = A^H$, the eigenvectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ can be assumed to be orthonormal. Call the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.
- Let $\vec{x} \in \mathbb{C}^n$ be arbitrary. Then $\vec{x} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_n\vec{v}_n$, where the coefficients are given by $\alpha_i = \langle \vec{x}, \vec{v}_i \rangle$ for $i = 1, 2, \ldots, n$.
- This implies that the constraint term obeys

$$\|\vec{x}\|_2^2 = \vec{x}^H\vec{x} = \left(\sum_{i=1}^{n} \alpha_i^* \vec{v}_i^H\right)\left(\sum_{j=1}^{n} \alpha_j \vec{v}_j\right) = \sum_{i=1}^{n} |\alpha_i|^2$$

and that the objective function obeys

$$\|\vec{x}\|_A^2 = \vec{x}^H A\vec{x} = \left(\sum_{i=1}^{n} \alpha_i^* \vec{v}_i^H\right) A \left(\sum_{j=1}^{n} \alpha_j \vec{v}_j\right) = \left(\sum_{i=1}^{n} \alpha_i^* \vec{v}_i^H\right) \cdot \left(\sum_{j=1}^{n} \alpha_j \lambda_j \vec{v}_j\right)$$

$A\vec{v}_j = \lambda_j \vec{v}_j$

$= \sum_{i=1}^{n} |\alpha_i|^2 \lambda_i$

Positive definite matrices: Geometry - 3

- Continuing, the objective function obeys

$$\|\vec{x}\|_A^2 = \vec{x}^H A \vec{x} = \sum_{i=1}^{n} |\alpha_i|^2 \lambda_i$$

.

$\max_{\vec{x}} \vec{x}^H A \vec{x}$ s.t. $\|\vec{x}\|_2^2 = 1.$

- Thus, the optimization problem becomes

$$\max_{\vec{\alpha}} \sum_{i=1}^{n} |\alpha_i|^2 \lambda_i \text{ such that } \sum_{i=1}^{n} |\alpha_i|^2 = 1.$$

- Suppose, without loss of generality, that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$.
- Then the optimal choice could be $\alpha_1 = 1$ and $\alpha_2 = \alpha_3 = \ldots = \alpha_n = 0$.
- Therefore, the maximum value attained equals λ_1 (i.e., the maximum eigenvalue), and the $\vec{x}$ that maximizes this objective function simply equals $\vec{v}_1$ (i.e., the eigenvector corresponding to the maximum eigenvalue).

$\vec{x} = \vec{v}_1.$

Positive definite matrices: Connection to matrix norms

Recall the 2-norm of a matrix A:

spectral norm.

$$\|A\|_2 = \sup_{\substack{\vec{x}\in\mathbb{C}^n \\ \|\vec{x}\|_2=1}} \|A\vec{x}\|_2.$$

Well, note that

$$\|A\vec{x}\|_2 = \sqrt{(A\vec{x})^H A\vec{x}} = \sqrt{\vec{x}^H A^H A\vec{x}} = \|\vec{x}\|_{A^H A}.$$

$\vec{x}^H(A^HA)\vec{x} = (A\vec{x})^H(A\vec{x}) = \|A\vec{x}\|_2^2 \geq 0.$

For any matrix A, it turns out that $A^H A$ is positive semi-definite. Thus, it follows that

A^HA is PSD.

$$\|A\|_2 = \sup_{\|\vec{x}\|_2=1} \|\vec{x}\|_{A^H A} = \sqrt{\lambda_{\max}(A^HA)}.$$

Positive definite matrices: Connection to matrix norms - 2

$A^3 = T\Lambda^3 T^H$

$A^n \stackrel{?}{=} T\Lambda^n T^H$.

Now, consider the special case where $A = A^H$.

- In this case,

$$\|\vec{x}\|_{\underbrace{A^H A}_{=A^2}} = \|\vec{x}\|_{A^2}.$$

- Also, since $A = T\Lambda T^H$ with T unitary, then $A^2 = T\Lambda^2 T^H$, and so $(\lambda_i(A))^2 = \lambda_i(A^2)$. Thus, when $A = A^H$ we have

$$\|A\|_2 = \sup_{\|\vec{x}\|_2=1} \|\vec{x}\|_{A^2} = \sqrt{\lambda_{\max}(A^2)} = \max_i |\lambda_i(A)|$$

- Similarly, it follows that $\|A^{-1}\|_2 = \dfrac{1}{\min_i |\lambda_i(A)|}$.

$A = T\Lambda T^H \Rightarrow A^{-1} = T\Lambda^{-1} T^H \Rightarrow \lambda_i(A^{-1}) = \dfrac{1}{\lambda_i(A)}$

$\|A^{-1}\|_2 = \max_i |\lambda_i(A^{-1})| = \max_i \dfrac{1}{|\lambda_i(A)|} = \dfrac{1}{\min_i |\lambda_i(A)|}$