MA-110 Linear Algebra and Differential Equations

Rekha Santhanam

Department of Mathematics
Indian Institute of Technology Bombay
Powai, Mumbai - 76

February 17, 2024
Lecture 19 D3

Diagonalizability: Summary

Thus: If an $n \times n$ matrix A has n linearly independent eigenvectors $v_1, \ldots, v_n$, then A is diagonalizable. Moreover, if $\lambda_1, \ldots, \lambda_n$ are the corresponding eigenvalues, then $P^{-1}AP = \Lambda$, where the diagonalizing matrix is $P = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$, and the diagonal matrix is $\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, i.e., $\boxed{P^{-1}AP = \Lambda}$, where

$\boxed{\text{The diagonal entries of } \Lambda \text{ are eigenvalues of } A \text{ and}}$

$\boxed{\text{The columns of } P \text{ are corresponding eigenvectors of } A.}$

$\boxed{\text{The EVD of } A \text{ is } A = P\Lambda P^{-1}.}$

Note: P need not be unique, e.g., replace v_1 by $2v_1$, etc.

Extra Reading: Simultaneous Diagonalizability

Assume $\boxed{A \text{ and } B \text{ are diagonalizable}}$. Then A and B have same eigenvector matrix S if and only if $AB = BA$.

Proof. ($\Rightarrow$) Assume $S^{-1}AS = \Lambda_1$ and $S^{-1}BS = \Lambda_2$, where Λ_1 and Λ_2 are diagonal matrices. Then

$AB = (S\Lambda_1 S^{-1})(S\Lambda_2 S^{-1}) = S(\Lambda_1\Lambda_2)S^{-1}$ and $BA = S(\Lambda_2\Lambda_1)S^{-1}$. Since $\Lambda_1\Lambda_2 = \Lambda_2\Lambda_1$, we get $AB = BA$.

(*Part of* $\Leftarrow$) Assume $AB = BA$. If $Ax = \lambda x$, then $ABx = B(Ax) = B(\lambda x) = \lambda Bx$. If $Bx = 0$, then x is an eigenvector of B, associated to $\mu = 0$.If $Bx \neq 0$, then x and Bx both are eigenvectors of A, associated to λ.

Special case: Assume all the eigenspaces of A are one dimensional. Then $Bx = \mu x$ for some scalar $\mu \Rightarrow x$ is an eigenvector of B. We will not prove the general case.

Eigenvalues of A^k

- If $Av = \lambda v$, then $A^2 v = A(Av) = A(\lambda v) = \lambda(Av) = \lambda^2 v$. Similarly $A^k v = \lambda^k v$ for any $k \geq 0$.

Thus if v is an eigenvector of A with associated eigenvalue λ, then v is also an eigenvector of A^k with associated eigenvalue λ^k for $k \geq 0$. If A is invertible, then $\lambda \neq 0$. Hence, the same also holds for $k < 0$ since $A^{-1} v = \lambda^{-1} v$.

- If A is diagonalizable , then $P^{-1}AP = \Lambda$ is diagonal where columns of P are eigenvectors of A.

Since $(P^{-1}A^k P) = \Lambda^k$, which is diagonal, we see that A^k is diagonalizable, and the eigenvectors of A^k are same as eigenvectors of A. Similarly, the same also holds for $k < 0$ if A is invertible.

Question: What is the EVD of A^k.

Let $F_0 = 0$, $F_1 = 1$ and $F_k = F_{k-1} + F_{k-2}$ for $k \geq 2$ define the Fibonacci sequence. What is the kth term?

If $u_k = \begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix}$, then $\begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_k \\ F_{k-1} \end{pmatrix}$, i.e.,

$u_k = Au_{k-1}$ for $n \geq 1$, where $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow u_k = A^k u_0$ for $k \geq 1$.

Characteristic polynomial of A: $\lambda^2 - \lambda - 1$; Eigenvalues:

$\lambda_1 = \dfrac{1+\sqrt{5}}{2}$, $\lambda_2 = \dfrac{1-\sqrt{5}}{2}$.

There are 2 distinct eigenvalues $\Rightarrow$ the associated eigenvectors x_1 and x_2 are linearly independent $\Rightarrow \{x_1, x_2\}$ is a basis for $\mathbb{R}^2$.

Write $u_0 = c_1x_1 + c_2x_2$. Then $u_k = A^k u_0 = A^k(c_1x_1 + c_2x_2)$

$$= c_1A^kx_1 + c_2A^kx_2 = c_1\left(\frac{1+\sqrt{5}}{2}\right)^k x_1 + c_2\left(\frac{1-\sqrt{5}}{2}\right)^k x_2.$$

Q: Find x_1, x_2, c_1 and c_2 and get the exact formula for F_k.

An Application: Predator-Prey Model

Let the owl and rat populations at time k be O_k and R_k respectively. Owls prey on the rats, so if there are no rats, the population of owls will go down by 50%. If there are no owls to prey on the rats, then the rat population will increase by 10%.

In particular, the rat and owl populations dependence is as follows:

$$\begin{aligned} O_{k+1} &= 0.5 O_k + 0.4 R_k \\ R_{k+1} &= -p O_k + 1.1 R_k \end{aligned}$$

The term $-p$ calculates the rats preyed by the owls.

Thus, if $P_k = \begin{pmatrix} O_k \\ R_k \end{pmatrix}$ and $A = \begin{pmatrix} 0.5 & 0.4 \\ -p & 1.1 \end{pmatrix}$, then $P_{k+1} = AP_k$ for all k. In particular, $P_k = A^k P_0$.

Exercise: If we start with a certain initial population of owls and rats, how many will be there in, say, 50 years, i.e., given P_0, what is P_{50}? What is the steady state, i.e., what is $\lim_{k\to\infty} P_k$?

An Application: Steady State

Suppose we have a system where the current state u_k depends on the previous one u_{k-1} linearly, i.e., $u_k = Au_{k-1}$. Then observe that $u_k = A^k u_0$. The steady state of the system is $u_\infty = \lim_{k\to\infty}(u_k)$. How do we find this?

- If u_0 is an eigenvector of A associated to λ, then $u_k = \lambda^k u_0$.
- Let $v_1, \ldots, v_r$ be eigenvectors of A associated respectively to $\lambda_1, \ldots, \lambda_r$. If $u_0 \in \mathrm{Span}\{v_1, \ldots, v_r\}$, i.e., $u_0 = c_1 v_1 + \cdots + c_r v_r$ for scalars $c_1, \ldots, c_r$, then $u_k = A^k u_0 = c_1 A^k v_1 + \cdots + c_r A^k v_r = c_1 \lambda_1^k v_1 + \cdots + c_r \lambda_r^k v_r$. In particular, if A is diagonalizable, then there is a basis of $\mathbb{R}^n$ of eigenvectors of A. Hence, this is applicable to every $u_0 \in \mathbb{R}^n$.

Let A be diagonalizable, and u_k represent population.

- Under what conditions will there be a population explosion?
- What conditions will force the population to become extinct?
- When does it stabilise (to a non-zero value)?

Hint: Depends on $|\lambda_i|$.

Extra Reading: Complex Eigenvalues

Example: Rotation by 90^o in $\mathbb{R}^2$ is given by $Q = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. It has no real eigenvectors since rotation by 90^o changes the direction.

Q has eigenvalues, but they are not real. $\det(Q - \lambda I) = \lambda^2 + 1$ $\Rightarrow \lambda_1 = i$ and $\lambda_2 = -i$, where $i^2 = -1$. Let us compute the eigenvectors.

$$(Q - iI)x_1 = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} x_1 = 0 \Rightarrow x_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}, \ (Q + iI)x_2 = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} x_2 = 0 \Rightarrow x_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

The eigenvalues, though imaginary, are distinct, hence eigenvectors are linearly independent.

If $P = \begin{pmatrix} x_1 & x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$, then $P^{-1}QP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

Conclusion: We need complex numbers $\mathbb{C}$ even if we are working with real matrices. Over $\mathbb{C}$, an $n \times n$ matrix A always has n eigenvalues.

Reason: $\boxed{\text{Fundamental theorem of Algebra}}$

Every polynomial over $\mathbb{C}$ of degree n has n roots in $\mathbb{C}$.

Inner product on $\mathbb{R}^n$

Defn. Define the **inner product** (dot product) of two vectors $v, w \in \mathbb{R}^n$ as $\boxed{v \cdot w = v^T w}$

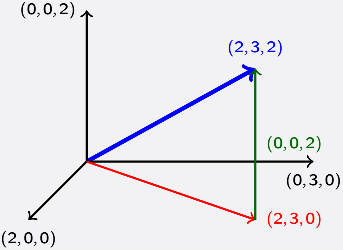
For v, w in $\mathbb{R}^n$ and c in $\mathbb{R}$

- $v \cdot w = v^T w = v_1 w_1 + \cdots + v_n w_n = w^T v = w \cdot v.$
- (Bilinearity)
$$\begin{aligned}(v+w)\cdot z &= (v+w)^T z = v^T z + w^T z = v \cdot z + w \cdot z \\ cv \cdot w &= (c\,v)^T w = c\,(v^T w) = v^T (c\,w) = v \cdot cw.\end{aligned}$$
- $v \cdot v = v^T v \geq 0$ and $v^T v = 0$ if and only if $v = 0$.

Define **length** (or norm) of v in $\mathbb{R}^n$ to be $\boxed{\|v\| = \sqrt{v \cdot v}.}$

Henceforth we will use $v^T w$ directly to write the dot product.

Let $v = (2,3,2)$. By Pythagoras theorem, $||v||$
$= \sqrt{||(2,3,0)||^2 + ||(0,0,2)||^2}$



$= \sqrt{2^2 + 3^2 + 2^2} = \sqrt{17}$.

Generalize by induction: Let $v = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$. Define
$||v|| = \sqrt{||(x_1, \cdots, x_{n-1}, 0)||^2 + ||(0, 0, \cdots, x_n)||^2}$
$= \sqrt{x_1^2 + \cdots + x_{n-1}^2 + x_n^2} = \sqrt{v^T v}$.

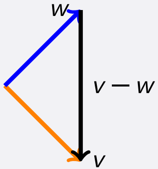
The length in $\mathbb{R}^n$ is compatible with the vector space structure.

Let $v, w \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then,

- $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$.
- $\|cv\| = |c|\|v\|$ • $\|v + w\| \leq \|v\| + \|w\|$ (Triangle Inequality)

Orthogonal vectors in $\mathbb{R}^n$

We say vectors v and w in $\mathbb{R}^n$ are orthogonal (perpendicular) if they satisfy the Pythagoras theorem, that is,
$||v||^2 + ||w||^2 = ||v - w||^2$



$$\begin{aligned}||v||^2 + ||w||^2 &= (v-w)^T(v-w)\\ &= (v^T - w^T)(v-w)\\ &= v^Tv - w^Tv + v^Tw + w^Tw\\ &= ||v||^2 - 2\ v^Tw + ||w||^2 \qquad (\text{since } w^Tv = v^Tw\ )\end{aligned}$$

Therefore, v and w are defined to be *orthogonal* if and only if $v^Tw = 0$.

Think! What can be said about Span$\{v\}$ and Span$\{w\}$ when v and w are orthogonal to each other in $\mathbb{R}^3$?

Orthogonal and Orthonormal Sets

Defn. A set of *non-zero* vectors $\{v_1, \cdots, v_k\} \subset \mathbb{R}^n$, is said to be an **orthogonal set** if $v_i^T v_j = 0$ **for all** $i, j = 1, \cdots n,\ i \neq j$.

Examples: $\{(1,3,1), (-1,0,1)\} \subset \mathbb{R}^3$,
$\{(2,1,0,-1), (0,1,0,1), (-1,1,0,-1)\} \subseteq \mathbb{R}^4$,
$\left\{(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}})\right\} \subseteq \mathbb{R}^3$, $\{e_1, \cdots, e_n\} \subseteq \mathbb{R}^n$.

Of these, the last two examples have all unit vectors (vectors of length one).

Defn. An orthogonal set $\{v_1, \cdots, v_k\} \subseteq \mathbb{R}^n$ with all unit vectors, i.e., $\|v_i\| = 1$ for all i, is called an **orthonormal set.**

Note: If $\{v_1, \cdots, v_k\}$ is an orthogonal set, then $\{u_1, \cdots, u_k\}$ is orthonormal, for $u_i = v_i/||v_i||$.

Exercise: If $S = \{v_1, \ldots, v_k\}$ is an orthogonal set, then v_k is orthogonal to each $v \in \text{Span}\{v_1, \ldots, v_{k-1}\}$.

Orthogonality and Linear Independence

Theorem: An orthogonal set in $\mathbb{R}^n$ is linearly independent.

Proof. Let $\{v_1, \cdots, v_k\}$ be an orthogonal set in $\mathbb{R}^n$, i.e. $v_i \neq 0$ and $v_i^T v_j = 0$ for $i \neq j$. Note that for $i = j$, $v_i^T v_i = ||v_i||^2 \neq 0$. Assume for some $a_1, \cdots, a_k \in \mathbb{R}$,

$$\begin{aligned}
a_1 v_1 + a_2 v_2 + \cdots + a_k v_k &= 0 \\
\Rightarrow (a_1 v_1 + a_2 v_2 + \cdots + a_k v_k)^T v_1 &= 0 \cdot v_1 = 0 \\
\Rightarrow (a_1 v_1^T + a_2 v_2^T + \cdots + a_k v_k^T)\ v_1 &= 0 \\
\Rightarrow a_1 v_1^T v_1 + a_2 v_2^T v_1 + \cdots + a_k v_k^T v_1 &= 0 \\
\Rightarrow a_1 ||v_1||^2 &= 0 \\
\Rightarrow a_1 &= 0 \text{ since } v_1 \neq 0
\end{aligned}$$

Similarly, we get $a_2 = \cdots = a_n = 0$. Hence $\{v_1, \cdots, v_k\}$ is linearly independent.

True/False: Any matrix whose columns form an orthogonal set is invertible. Give example

Matrices with Orthogonal Columns

Let $A = [v_1 \ \cdots \ v_n]$ be $m \times n$. If $\{v_1, \ldots, v_n\}$ form an *orthonormal* set in $\mathbb{R}^m$, then

$$A^T A = \begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix} = \begin{pmatrix} v_1^T v_1 & \ldots & v_1^T v_n \\ \vdots & & \vdots \\ v_n^T v_1 & \ldots & v_n^T v_n \end{pmatrix} = I_n.$$

Defn. A square matrix A whose column vectors form an orthonormal set is called an orthogonal matrix.

If $Q = [u_1 \ \cdots \ u_n]$ is an orthogonal matrix, then

- $\{u_1, \ldots, u_n\}$ is an orthonormal set (by definition)
- $Q^T Q = I = QQ^T$ Why?
- $\|Qv\| = \sqrt{(Qv)^T (Qv)} = \sqrt{v^T Q^T Q v} = \sqrt{v^T x} = \|v\|.$

$\Rightarrow$ the only (real) eigenvalues of Q, if they exist, are ± 1.

- Row vectors of Q are orthonormal since $QQ^T = I$.

Orthogonal Matrices: Examples

Examples: 1. $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. 2. $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

3. $\frac{1}{3}\begin{pmatrix} 2 & 1 & -2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix}$. 4. $\frac{1}{2}\begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{pmatrix}$.