MA-106 Linear Algebra

H. Ananthnarayan

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

15th February 2018
D1 - Lecture 19

Summary: Diagonalizability

Let A be $n \times n$.

• A is diagonalizable $\Leftrightarrow \mathbb{R}^n$ has a basis, say $\mathcal{B} = \{v_1, \ldots, v_n\}$, of eigenvectors of A, associated to eigenvalues $\lambda_1, \ldots, \lambda_n$. In this case, $P^{-1}AP = \Lambda$, where $P = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ and Λ is a diagonal matrix with entries $\lambda_1, \ldots, \lambda_n$.

• If A is diagonalizable, and T is the linear operator defined by $Tx = Ax$, then $[T]_{\mathcal{B}}^{\mathcal{B}} = \Lambda$. Thus diagonalization of A is the same as finding a basis w.r.t. which the matrix of T (defined by $Tx = Ax$) is diagonal.

• Eigenvectors associated to distinct eigenvalues are linearly independent. In particular, if A has n distinct eigenvalues, A is diagonalizable.

• If v is an eigenvector of A with respect to eigenvalue λ, then v is also an eigenvector of A^k w.r.t. eigenvalue λ^k for $k \geq 0$. In particular, if A is diagonalizable, and $P^{-1}AP$ is diagonal, the same holds for $P^{-1}A^kP$. These statements hold for all $k \in \mathbb{Z}$ if A is invertible.

Complex Eigenvalues

Ex: Rotation by 90^o in $\mathbb{R}^2$ is given by $Q = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

$Qe_1 =$ __ and $Qe_2 =$ __.It has no eigenvectors since rotation by 90^o changes the direction. It has no real eigenvectors.

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 $P^{-1}QP = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

Complex Vectors

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 $\mathbb{C}$.

For a complex number $x = a + ib$, its conjugate is $\bar{x} = a - ib$ and $|x|^2 = \bar{x} \cdot x = a^2 + b^2 \in \mathbb{R}$ is the length (or modulus) of x.
If $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{C}^n$ are complex vectors,
define $\boxed{x \cdot y = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n} \in \mathbb{C}$. Then

- $x \cdot x = |x_1|^2 + \cdots + |x_n|^2 \in \mathbb{R}$ and is ≥ 0.
- $x \cdot x = 0 \Leftrightarrow x = 0$.
- For $c \in \mathbb{C}$, $x \cdot cx = \bar{x}_1 c x_1 + \cdots + \bar{x}_n c x_n = (x \cdot x)c$.

This defines a *complex inner product* (or dot product) on $\mathbb{C}^n$.

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

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)

$$(v + w) \cdot z = (v + w)^T z = v^T z + w^T z = v \cdot z + w \cdot z$$
$$cv \cdot w = (c\, v)^T w = c\,(v^T w) = v^T (c\, w) = v \cdot cw.$$

- $v \cdot v = v^T v \geq 0$ and $v^T v = 0$ if and only if $v = 0$.

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

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)

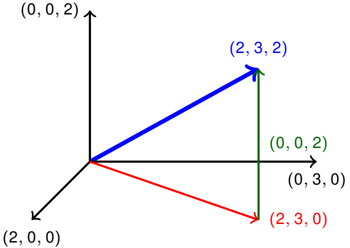
Henceforth we will use $v^T w$ directly to write the dot product.

Reading : Length of a vector in $\mathbb{R}^3$ and $\mathbb{R}^n$

Let $v = (2,3,2)$. By Pythagoras theorem, $||v||$

$= \sqrt{||(2,3,0)||^2 + ||(0,0,2)||^2}$

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



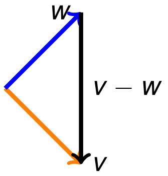
Generalize by induction: Let $v = (x_1, \cdots, x_n)^T \in \mathbb{R}^n$. Define

$$\|v\| = \sqrt{\|(x_1, \cdots, x_{n-1}, 0)\|^2 + \|(0, 0, \cdots, x_n)\|^2} = \sqrt{x_1^2 + \cdots + x_{n-1}^2 + x_n^2} = \sqrt{v^T v}.$$

The length (or norm) of a vector 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) $\Leftrightarrow$ they satisfy the Pythagoras theorem $\Leftrightarrow$ $||v||^2 + ||w||^2 = ||v - w||^2$

$$\begin{aligned}
\Leftrightarrow \|v\|^2 + \|w\|^2 &= (v - w)^T(v - w) \\
&= (v^T - w^T)(v - w) \\
&= v^T v - w^T v - v^T w + w^T w \\
&= \|v\|^2 - 2\, v^T w + \|w\|^2 \quad (\text{since } w^T v = v^T w\,)
\end{aligned}$$

Therefore, v and w are said to be *orthogonal* if and only if

$$v^T w = 0.$$

Q: 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 $u_i = v_i/||v_i||$.

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 $a_1, \cdots, a_k \in \mathbb{R}$,

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

Similarly, we can show that $a_2 = \cdots = a_n = 0$.
Hence the set $\{v_1, \cdots, v_k\}$ is linearly independent.

Matrices with Orthogonal Columns

True/False: Any matrix whose columns form an orthogonal set is invertible. Why?

$$A = \begin{pmatrix} 2 & 1 & -2 \\ 2 & -2 & 1 \\ 1 & 2 & 2 \end{pmatrix}. \text{ and } B = \begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{pmatrix} \text{ are both}$$

examples of such matrices!

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

Orthogonal Matrices

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?
- $\|Qx\| = \sqrt{(Qx)^T(Qx)} = \sqrt{x^T Q^T Qx} = \sqrt{x^T x} = \|x\|$.
- Row vectors of Q are orthonormal (since $QQ^T = I$).

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