Chapter 0 Preliminaries

Objective of the course

Introduce basic matrix results and techniques that are useful in theory and applications.

It is important to know many results, but there is no way I can tell you all of them.

I would like to introduce techniques so that you can obtain the results you want.

This is the standard teaching philosophy of

"It is better to teach students how to fish instead of just giving them fishes."

Assumption You are familiar with

- Linear equations, solution sets, elementary row operations.
- Matrices, column space, row space, null space, ranks,
- Determinant, eigenvalues, eigenvectors, diagonal form.
- Vector spaces, basis, change of bases.
- Linear transformations, range space, kernel.
- Inner product, Gram Schmidt process for real matrices.
- Basic operations of complex numbers.

$A\vec{x}=\vec{b}$ $\quad [A|b]$

$A=\underbrace{[a_1|a_2|\cdots|a_n]}_{n}\}m$

A $n\times n$

$x\neq 0$

$Ax=\lambda x$

$\det(A-\lambda I)=0$

Notation

- $M_n(\mathbb{F})$, $M_{m,n}(\mathbb{F})$ are the set of $n\times n$ and $m\times n$ matrices over $\mathbb{F}=\mathbb{R}$ or $\mathbb{C}$ (or a more general field).
- $\mathbb{F}^n$ is the set of column vectors of length n with entries in $\mathbb{F}$.
- Sometimes we write $M_n, M_{m,n}$ if $\mathbb{F}=\mathbb{C}$.
- If $A=\begin{pmatrix} A_1 & 0 \\ 0 & A_2\end{pmatrix}\in M_{m+n}(\mathbb{F})$ with $A_1\in M_m(\mathbb{F})$ and $A_2\in M_n(\mathbb{F})$, we write $A=A_1\oplus A_2$.
- x^t, A^t denote the transpose of a vector x and a matrix A.
- For a complex matrix A, $\overline{A}$ denotes the matrix obtained from A by replacing each entry by its complex conjugate. Furthermore, $A^*=(\overline{A})^t$.

Linear Transformations. $\boxed{T:\mathbb{R}^m\to\mathbb{R}^n}$

$T: V_1\to V_2$

$T(x+y)=T(x)+T(y)$

$T(\lambda x)=\lambda T(x)$

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

$Col(A)\subseteq\mathbb{R}^3$

basis $=\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\1\\0\end{bmatrix}\right\}$

$v=u_1\begin{bmatrix}1\\0\\0\end{bmatrix}+u_2\begin{bmatrix}2\\1\\0\end{bmatrix}+\cancel{u_3\begin{bmatrix}1\\2\\0\end{bmatrix}}$

$\mathrm{Ker}(T) = \{ x \in V_1 : T(x) = 0 \} \subseteq V_1$

$\mathrm{Range}(T) = \{ T(x) \in V_2 : x \in V_1 \} \subseteq V_2$

~~If $T(x) = Ax$~~

If $T : \mathbb{R}^n \to \mathbb{R}^m$ ~~$y = Ax$~~

and $T(x) = Ax$ for an $m \times n$ A.

then $\mathrm{Ker}(T) = \{ x : Ax = 0 \} = \mathrm{Null}(A)$

$\mathrm{Range}(T) = \{ Ax : x \in \mathbb{R}^n \} = \mathrm{Col}(A)$.

Inner product of $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$

$$x \cdot y = x_1 y_1 + x_2 y_2 + x_3 y_3$$

$$\|x\| = (x \cdot x)^{1/2} = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

If $x \cdot y = 0$ then x and y are orthogonal/perpendicular

$\boxed{\text{Notation:}}$ $\underset{2\times 2}{A_1} \oplus \underset{3\times 3}{A_2}$

$$\begin{array}{cc} & 2 \\ 2 & \left[\begin{array}{c|c} A_1 & 0 \\ \hline & A_2 \end{array}\right] \end{array} \begin{array}{l} - 2\times 3 \\ - 3\times 3 \end{array}$$

A

$A = \begin{bmatrix} 2 & 1-i & i \\ 2 & 3 & 2+3i \end{bmatrix}$, $\bar{A} = \begin{bmatrix} 2 & 1+i & -i \\ 2 & 3 & 2-3i \end{bmatrix}$

$$A^* = (\bar{A})^t = \begin{bmatrix} 2 & 2 \\ 1+i & 3 \\ -i & 2-3i \end{bmatrix}$$

1 Similarity, Jordan form, and applications

1.1 Similarity

Definition 1.1 *Two matrices $A, B \in M_n$ ~~are similar~~ if there is an invertible $S \in M_n$ such that $S^{-1}AS = B$, equivalently, $\boxed{A = T^{-1}BT \text{ with } T = S^{-1}}$.*

Remark 1.2 *If A and B are similar, then $\det(A) = \det(B)$ and $\det(zI - A) = \det(zI - B)$. Furthermore, if $S^{-1}AS = B$, then $Bx = \mu x$ if and only if $A(\underbrace{Sx}_{y}) = \mu(\underbrace{Sx}_{y})$.*

$Ay = \lambda y$

$\Rightarrow B(Ty) = \lambda(Ty) = B(S^{-1}y)$

$= \lambda(S^{-1}y)$

Remark: Similarity is an equivalence relation on $n \times n$ matrices

(R) $\forall A \in M_n$, $A = IAI$ $\therefore$ reflexive

(S) $S^{-1}AS = B \Rightarrow A = TBT^{-1}$ with $T = S^{-1}$ $\therefore$ Symmetric

(T) $A \sim B$, $B \sim C$ $\quad B = S^{-1}AS$, $C = \hat{S}^{-1}B\hat{S}$.

Then $\quad C = \hat{S}^{-1}B\hat{S} = \hat{S}^{-1}(S^{-1}AS)\hat{S}$

$= (S\hat{S})^{-1}A(S\hat{S})$

Remark 1.2 Proof

If $B = S^{-1}AS$, then $\det(B) = \det(S^{-1}AS) = \det(S^{-1})\det(A)\det(S)$

$= \det(A)\det(S)\det(S^{-1})$

$= \det(A)\det(SS^{-1}) = \det(A)\det(I)$

$= \det(A)$.

If $B = S^{-1}AS$, then $\det(zI - B) = \det(zI - S^{-1}AS)$

$= \det(z(S^{-1}S) - S^{-1}AS)$

$= \det(S^{-1}(zI)S - S^{-1}AS)$

$= \det(S^{-1}(zI - A)S)$

$= \det(zI - A)$

i.e, $\exists$ invertible S s.t. $S^{-1}AS = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$

Theorem 1.3 *Let $A \in M_n(\mathbb{F})$. Then A is diagonalizable over $\mathbb{F}$ if and only if it has n linearly independent eigenvectors in $\mathbb{F}^n$.*

Reason:
$$S^{-1}AS = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

$$\Leftrightarrow \quad \underset{n\times n}{A} [x_1 | \cdots | x_n] = \underbrace{[x_1 | \cdots | x_n]}_{S} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

$$\Leftrightarrow \quad [Ax_1 | Ax_2 | \cdots | Ax_n] = [\lambda_1 x_1 | \lambda_2 x_2 | \cdots | \lambda_n x_n]$$

$$\Leftrightarrow \quad Ax_i = \lambda_i x_i \ \ \forall i = 1, \ldots, n. \quad \& \quad (x_1 \cdots x_n) \text{ is invertible, i.e. } \{x_1 \cdots x_n\} \text{ is linearly independent.}$$

Remark 1.4 *Suppose $A = SDS^{-1}$ with $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Let S have columns $x_1, \ldots, x_n$ and S^{-1} have rows $y_1^t, \ldots, y_j^t$. Then*

$$Ax_j = \lambda_j x_j \quad \text{and} \quad y_j^t A = \lambda_j y_j^t, \qquad j = 1, \ldots, n,$$

and

$$\boxed{A = \sum_{j=1}^{n} \lambda_j x_j y_j^t.}$$

Thus, $x_1, \ldots, x_n$ are the (right) eigenvectors of A, and $y_1^t, \ldots, y_n^t$ are the left eigenvectors of A.

$$S^{-1}A = DS^{-1}$$

$$\Leftrightarrow \quad \begin{bmatrix} y_1^t \\ \vdots \\ y_n^t \end{bmatrix} A = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} y_1^t \\ \vdots \\ y_n^t \end{bmatrix}$$

$$\Leftrightarrow \quad \begin{bmatrix} y_1^t A \\ \vdots \\ y_n^t A \end{bmatrix} = \begin{bmatrix} \lambda_1 y_1^t \\ \vdots \\ \lambda_n y_n^t \end{bmatrix}$$

$$A = SDS^{-1} = [x_1 | \cdots | x_n] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} y_1^t \\ \vdots \\ y_n^t \end{bmatrix}$$

$$= [\lambda_1 x_1 | \cdots | \lambda_n x_n] \begin{bmatrix} y_1^t \\ \vdots \\ y_n^t \end{bmatrix}$$

$$= \lambda_1 x_1 y_1^t + \lambda_2 x_2 y_2^t + \cdots + \lambda_n x_n y_n^t.$$

Note if

$$\hat{A} = \mu_1 x_1 y_1^t + \cdots + \mu_n x_n y_n^t$$

$$= S \begin{bmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{bmatrix} S^{-1}$$

Remark 1.5 *For a real matrix $A \in M_n(\mathbb{R})$, we may not be able to find real eigenvalues, and we may not be able to find n linearly independent eigenvectors even if* $\det(zI - A) = 0$ *has only real roots.*

$-\det(\lambda I - A) = \lambda^2 + 1$

Example 1.6 *Consider* $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ *and* $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. *Then A has no real eigenvalues, and B does not have two linearly independent eigenvectors.*

Remark 1.7 *By the fundamental theorem of algebra, every complex polynomial can be written as a product of linear factors. Consequently, for every $A \in M_n(\mathbb{C})$, the characteristic polynomial has linear factors:*

$$\det(zI - A) = (z - \lambda_1) \cdots (z - \lambda_n).$$

So, we needs only to find n linearly independent eigenvectors.