Lecture #11 – Tuesday, February 10, 2004

Orthogonality

Definition

Let X, Y be vectors in $\mathbf{R}^n$. The <u>dot product</u> is $X \cdot Y = X^T Y \in \mathbf{R}$.

So, if $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $X \cdot Y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{j=1}^{n} x_j y_j$.

Definition

The <u>length</u> of a vector X, denoted $\|X\| = \sqrt{X \cdot X}$. So if $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\|X\| = \sqrt{{x_1}^2 + {x_2}^2 + \cdots + {x_n}^2} \geq 0$.

Example

Let $X = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 4 \end{pmatrix}$, $Y = \begin{pmatrix} -1 \\ 2 \\ 3 \\ -2 \end{pmatrix}$. Find $X \cdot Y$ and $\|X\|$.

- $X \cdot Y = (1)(-1) + (0)(2) + (-2)(3) + (4)(-2) = -15$.
- $\|X\| = \sqrt{(1)^2 + (0)^2 + (-2)^2 + (4)^2} = \sqrt{21}$.

Theorem

Let X, Y, Z be vectors in $\mathbf{R}^n$. Then:

1) $X \cdot Y = Y \cdot X$. (Proof: $\sum_{j=1}^{n} x_j y_j = \sum_{j=1}^{n} y_j x_j$)
2) $X \cdot (Y + Z) = X \cdot Y + X \cdot Z$.
3) For $a \in \mathbf{R}$, $(aX) \cdot Y = X \cdot (aY) = a(X \cdot Y)$.
4) $\|X\| \geq 0$, $\|X\| = 0 \Leftrightarrow x = 0$.

Example

Let X be a vector in $\mathbf{R}^n$, $X \neq 0$. Find all $Y \in \mathbf{R}^n$ such that Y is collinear to X and is unitary.

- Let $Y = aX, a \in \mathbf{R}$. $\|Y\| = \|aX\| = |a|\|X\|$. We want $\|Y\| = 1$, so $|a|\|X\| = 1 \Rightarrow |a| = \frac{1}{\|X\|} \Rightarrow a = \pm \frac{1}{\|X\|}$.
- So, $Y = \pm \frac{1}{\|X\|} X$.

Definition

Let X, Y be two vectors in $\mathbf{R}^n$. X and Y are <u>orthogonal</u> iff $X \cdot Y = 0$.
Let $\{X_1, X_2, \cdots, X_k\}$ be a set of vectors in $\mathbf{R}^n$. $\{X_1, X_2, \cdots, X_k\}$ is <u>orthogonal</u> iff $x_i \neq 0, i = 1,2,\cdots,k$ and $X_i X_j = 0, i \neq j$.

Moreover, if $\|X_j\| = 1$, $j = 1,2,\cdots,k$, then $\{X_1, X_2, \ldots, X_k\}$ is <u>orthonormal</u>.

Example

The standard basis of $\mathbf{R}^n$ is a orthonormal set. $E_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $E_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $E_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Definition

If $\{X_1, X_2, \cdots, X_k\}$ is an <u>orthogonal set</u>, then $\{a_1 X_1, a_2 X_2, \cdots, a_k X_k\}$ is also an <u>orthogonal set</u> if $a_j \neq 0, j = 1,2,\cdots,k$.

For $a_j = \dfrac{1}{\|X_j\|}$, the set $\{a_1 X_1, a_2 X_2, \cdots, a_k X_k\}$ will be <u>orthonormal</u>.

Example

Let $X_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$, $X_2 = \begin{pmatrix} 0 \\ -3 \\ 3 \end{pmatrix}$, $X_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$. Construct an orthonormal set.

- $\{X_1, X_2, X_3\}$ is an orthogonal set because:
  - $X_i \neq 0$.
  - $X_1 X_2 = 0$, $X_1 X_3 = 0$, $X_2 X_3 = 0$.
- Find $a_j, j = 1,2,3$ such that $\{a_1 X_1, a_2 X_2, a_3 X_3\}$ is orthonormal:

  $a_1 = \dfrac{1}{\|X_1\|} = \dfrac{1}{\sqrt{(2)^2 + (1)^2 + (1)^2}} = \dfrac{1}{\sqrt{6}}$, $a_2 = \dfrac{1}{3\sqrt{2}}$, $a_3 = \dfrac{1}{\sqrt{3}}$.
- So, $\left\{ \dfrac{X_1}{\sqrt{6}}, \dfrac{X_2}{3\sqrt{2}}, \dfrac{X_3}{\sqrt{3}} \right\}$ is orthonormal.

Theorem (Pythagorean)

If X and Y are orthogonal, then $\|X + Y\|^2 = \|X\|^2 + \|Y\|^2$.

Proof

X and Y are orthogonal; it means $X \cdot Y = 0$. So,
$\|X + Y\|^2 = (X + Y) \cdot (X + Y) = X \cdot X + X \cdot Y + Y \cdot X + Y \cdot Y = X \cdot X + 0 + 0 + Y \cdot Y = \|X\|^2 + \|Y\|^2$.

Theorem

Every orthogonal set of vectors of $\mathbf{R}^n$ is <u>independent</u>.

Proof

Let $\{X_1, X_2, \cdots, X_k\}$ be an orthogonal set. Consider $Y = t_1 X_1 + t_2 X_2 + \cdots + t_k X_k = 0$. We want to prove that $t_j = 0$.

$$X_j \cdot Y = 0 \Rightarrow X_j \cdot (t_1 X_1 + t_2 X_2 + \cdots + t_k X_k) = t_1 X_1 \cdot X_j + t_2 X_2 \cdot X_j + \cdots + t_j X_j \cdot X_j + \cdots + t_k X_k = 0$$

$$\Rightarrow t_j \|X_j\|^2 = 0 \Rightarrow t_j = 0 .$$

Theorem

If $\{E_1, E_2, \cdots, E_n\}$ is an orthogonal basis of $\mathbf{R}^n$, then $\forall X \in \mathbf{R}^n$,

$$X = \frac{X \cdot E_1}{\|E_1\|^2} E_1 + \frac{X \cdot E_2}{\|E_2\|^2} E_2 + \cdots + \frac{X \cdot E_n}{\|E_n\|^2} E_n .$$

Proof

Let $X \in \mathbf{R}^n$. Then $X = t_1 E_1 + t_2 E_2 + \cdots + t_n E_n$ because is a $\{E_1, E_2, \cdots, E_n\}$ basis of $\mathbf{R}^n$.

$$X \cdot E_j = (t_1 E_1 + t_2 E_2 + \cdots t_n E_n) \cdot E_j = t_1 E_1 \cdot E_j + \cdots + t_j E_j \cdot E_j + \cdots + t_n E_n \cdot E_j = t_j \|E_j\|^2 \Rightarrow t_j = \frac{X \cdot E_j}{\|E_j\|^2} .$$

Example

Let $X_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $X_2 = \begin{pmatrix} 0 \\ -3 \\ 3 \end{pmatrix}$, $X_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$. Show that $\{X_1, X_2, X_3\}$ is an orthogonal basis of $\mathbf{R}^3$.

- It is enough to prove $X_1 \cdot X_2 = 0$, $X_2 \cdot X_3 = 0$, $X_1 \cdot X_3 = 0$ because:
    - $\{X_1, X_2, X_3\}$ is an orthogonal set, so X_1, X_2, X_3 are independent.
    - $\dim R^3 = 3$. So a set of 3 linearly independent vectors is a basis.

Example

Write $X = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ as a linear combination of $X_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $X_2 = \begin{pmatrix} 0 \\ -3 \\ 3 \end{pmatrix}$, $X_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.

- Since $\{X_1, X_2, X_3\}$ is orthogonal,

  $$X = \frac{X \cdot X_1}{\|X_1\|^2} X_1 + \frac{X \cdot X_2}{\|X_2\|^2} X_2 + \frac{X \cdot X_3}{\|X_3\|^2} X_3 = \frac{2a+b}{5} X_1 + \frac{-3b+c}{6} X_2 + \frac{a-b+c}{3} X_3 .$$

Theorem: Orthogonal Lemma

Let $\{E_1, E_2, \cdots, E_m\}$ be an orthogonal set in $\mathbf{R}^n$. For $X \in \mathbf{R}^n$,

$E_{m+1} = X - \left[\frac{X \cdot E_1}{\|E_1\|^2} E_1 + \frac{X \cdot E_2}{\|E_2\|^2} E_2 + \cdots + \frac{X \cdot E_m}{\|E_m\|^2} E_m\right]$. Then:

1) $E_{m+1} \cdot E_j = 0,\ j = 1,2,\cdots,m$.
2) If $E_{m+1} = 0$, then $X \in \operatorname{span}(E_1, E_2, \cdots, E_m)$.
3) If $E_{m+1} \neq 0$, then $\{E_1, E_2, \cdots, E_m, E_{m+1}\}$ is an orthogonal set.

Proof

- Suppose (1) is proven.
- If $E_{m+1} = 0$, then
  $$X - \left[\frac{X \cdot E_1}{\|E_1\|^2} E_1 + \frac{X \cdot E_2}{\|E_2\|^2} E_2 + \cdots + \frac{X \cdot E_m}{\|E_m\|^2} E_m\right] = 0 \Rightarrow X = \frac{X \cdot E_1}{\|E_1\|^2} E_1 + \frac{X \cdot E_2}{\|E_2\|^2} E_2 + \cdots + \frac{X \cdot E_m}{\|E_m\|^2} E_m .$$
  $\Rightarrow X \in \operatorname{span}(E_1, E_2, \cdots, E_m)$
- If $E_{m+1} \neq 0$, then (1) gives that $E_{m+1} \cdot E_j = 0,\ j = 1,2,\cdots,m$, and $X_i X_j = 0, i \neq j \leq m$. It means that $\{E_1, E_2, \cdots, E_m, E_{m+1}\}$ is an orthogonal set.
  $$\begin{aligned} E_{m+1} \cdot E_j &= \left(X - \left[\frac{X \cdot E_1}{\|E_1\|^2} E_1 + \frac{X \cdot E_2}{\|E_2\|^2} E_2 + \cdots + \frac{X \cdot E_m}{\|E_m\|^2} E_m\right]\right) \cdot E_j \\ &= X \cdot E_j - \frac{X \cdot E_1}{\|E_1\|^2} E_1 \cdot E_j - \frac{X \cdot E_2}{\|E_2\|^2} E_2 \cdot E_j - \cdots - \frac{X \cdot E_j}{\|E_j\|^2} E_j \cdot E_j - \cdots - \frac{X \cdot E_m}{\|E_m\|^2} E_m \cdot E_j . \\ &= X \cdot E_j - X \cdot E_j = 0 \end{aligned}$$

Theorem

Let U be a subspace of $\mathbf{R}^n$, and $\{X_1, X_2, \cdots, X_m\}$ is an orthogonal set of U. Then:

1) $\{X_1, X_2, \cdots, X_m\}$ can be extended in a orthogonal basis of U.
2) If $U \neq \{0\}$, then there are orthogonal basis.

Proof

1) Gram-Schmidt Orthogonalisation.
2) $U \neq \{0\}$. Let $X_1 \in U, X_1 \neq 0$. $\{X_1\}$ is orthogonal and (1) gives the result.

Gram-Schmidt Orthogonalisation

Let U be a subspace of $\mathbf{R}^n$. Let $\{X_1, X_2, \cdots, X_m\}$ be a basis of U. We'll construct $\{E_1, E_2, \cdots, E_m\}$ orthogonal:

- $E_1 = X_1$.
- $E_2 = X_2 - \dfrac{X_2 \cdot E_1}{\|E_1\|^2} E_1$.

- $$E_k = X_k - \left[\frac{X_k \cdot E_1}{\|E_1\|^2}E_1 + \frac{X_k \cdot E_2}{\|E_2\|^2}E_2 + \cdots + \frac{X_k \cdot E_{k-1}}{\|E_{k-1}\|^2}E_{k-1}\right] = X_k - \sum_{j=1}^{k-1}\frac{X_j \cdot E_j}{\|E_j\|^2}E_j, k \le m.$$

Lecture #12 – Thursday, February 12, 2004

PROJECTIONS

Definition

Let U be a subspace of $\mathbf{R}^n$. The orthogonal complement of U is denoted $U^\perp = \{X \in \mathbf{R}^n, X \cdot Y = 0 \text{ for any } Y \in U\}$.

Proposition

1) $U^\perp$ is a subspace of $\mathbf{R}^n$.
2) If $U = \text{span}\{X_1, X_2, \cdots, X_k\}$, then $U^\perp = \{X \in \mathbf{R}^n, X \cdot X_j = 0, j = 1,2,\cdots,k\}$.

Proof

1) $U^\perp$ is a subspace of $\mathbf{R}^n$.
   - $0 \in U^\perp$: $0 \cdot Y = 0, \forall Y \in U$.
   - If $X, Y \in U^T$, then $X + Y \in U^\perp$: $\forall Z \in U$, $X \cdot Z = 0$ and $Y \cdot Z = 0$. $(X+Y)\cdot Z = X \cdot Z + Y \cdot Z = 0$. So $X + Y \in U^\perp$.
   - $a \in \mathbf{R}, X \in U^\perp$, then $aX \in U^\perp$: $\forall Z \in U$, $X \cdot Z = 0$. $(aX)\cdot Z = a(X \cdot Z) = a(0) = 0$. So $aX \in U^T$.
2) If $U = \text{span}\{X_1, X_2, \cdots, X_k\}$, then $U^\perp = \{X \in \mathbf{R}^n, X \cdot X_j = 0, j = 1,2,\cdots,k\}$.
   - Let $V = \{X \in \mathbf{R}^n, X \cdot X_j = 0, j = 1,2,\cdots,k\}$, and $U^\perp = \{X \in \mathbf{R}^n, X \cdot Y = 0 \text{ for any } Y \in U\}$.
   - Then $U^\perp \subset V$: Let $X \subset U^\perp$. $X \cdot X_j = 0$ since $X_j \in U$. So $U^\perp \subset V$.
   - Also, $V \subset U^\perp$. Let $X \in V$ and $Y \in U = \text{span}\{X_1, X_2, \cdots, X_k\}$.
     $\begin{array}{l} Y = t_1X_1 + t_2X_2 + \cdots t_kX_k \Rightarrow X \cdot Y = X \cdot (t_1X_1 + t_2X_2 + \cdots t_kX_k) \\ = t_1(X_1 \cdot X) + t_2(X_2 \cdot X) + \cdots + t_k(X_1 \cdot X) = 0 \end{array}$. So $X \in U^\perp$.
   - So, $V = U^\perp$.

Example

Consider $U = \text{span}\left\{\begin{pmatrix}1\\-2\\3\end{pmatrix}, \begin{pmatrix}-1\\1\\1\end{pmatrix}\right\} \subset \mathbf{R}^3$. Find $U^\perp$.

- Let $X = \begin{pmatrix}x\\y\\z\end{pmatrix} \in U^\perp$. It means $X \cdot \begin{pmatrix}1\\-2\\3\end{pmatrix} = 0$ and $X \cdot \begin{pmatrix}-1\\1\\1\end{pmatrix} = 0$. So, $\begin{cases} x - 2y + 3z = 0 \\ -x + y + z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 5s \\ y = 4s, s \in \mathbf{R} \\ z = s \end{cases}$.

- $X=\begin{pmatrix}x\\y\\z\end{pmatrix}=s\begin{pmatrix}5\\4\\1\end{pmatrix}$. So, $U^{\perp}=\operatorname{span}\left\{\begin{pmatrix}5\\4\\1\end{pmatrix}\right\}$.

Proposition

Let $\{E_1, E_2, \cdots, E_m\}$ be an orthogonal basis of $U \in \mathbf{R}^n$. Define $\forall X \in \mathbf{R}^n$,
$P(X)=\frac{X\cdot E_1}{\|E_1\|^2}E_1+\frac{X\cdot E_1}{\|E_1\|^2}E_1+\cdots+\frac{X\cdot E_m}{\|E_m\|^2}E_m$. Then:

1) $P(X)\in U$ and $X-P(X)\in U^{\perp}$.
2) $P(X)$ is independent of the choice of the <u>orthogonal basis</u>. $P(X)$ is the <u>orthogonal projection</u> on *U*.
3) $P(X)$ satisfies $\|X-P(X)\|<\|X-Y\|$ for all $Y\in U$, $Y\neq P(X)$.
4) $\dim U+\dim U^{\perp}=\dim \mathbf{R}^n$.

Proof

1) Want $X-P(X)\in U^{\perp}$. Enough to prove that $P(X)\cdot E_j=0,\ j=1,2,\cdots,k$.
2) Consider another <u>orthogonal basis</u> of *U*. Let $P'=P'(X)$ with the new basis.
   $Y=P-P'=P-X-(P'-X)$. Since $P,P'\in U$, $P-P'\in U$. Also, $P-X\in U^{\perp}$ and $P'-X\in U^{\perp}$. So $Y\in U\cap U^{\perp}\Rightarrow Y\cdot Y=0\Rightarrow Y=0$. Since $Y=P-P'\Rightarrow P=P'$.

APPLICATION: $\mathbf{R}^3$

- Equation of a line through P_0 and with direction $\vec{v}\in\mathbf{R}^3$ is $L=\left\{P, \overrightarrow{PP_0}=t\vec{v}, t\in\mathbf{R}\right\}$ (parametric equation of *L*).
- Plane in $\mathbf{R}^3$ through P_0 with direction $\vec{u},\vec{v}\in\mathbf{R}^3$ independent is $P=\{P=P_0+t\vec{u}+s\vec{v}, t, s\in R\}$. Plane through P_0 orthogonal to $\vec{n}\in\mathbf{R}^3$ is $P=\left\{P, \overrightarrow{PP_0}\cdot\vec{n}=0\right\}$.
  - Let $P_0=(x_0, y_0, z_0)$, $\vec{n}=\begin{pmatrix}a\\b\\c\end{pmatrix}$. Then
    $P=\left\{P(x,y,z), \begin{pmatrix}x_0-x\\y_0-y\\z_0-z\end{pmatrix}\cdot\begin{pmatrix}a\\b\\c\end{pmatrix}=0\right\}=\{P(x,y,z), a(x_0-x)+b(y_0-y)+c(z_0-z)=0\}$ (scalar equation of a plane).

Example

Plane through $P_0(-1,2,-3)$ orthogonal to $\vec{n}=\begin{pmatrix}1\\2\\-3\end{pmatrix}$.

- $P(x,y,z)\in P$ if $\overrightarrow{P_0P}\cdot\vec{n}=0$. So $(1)(x+1)+(2)(y-2)+(-3)(z+3)=0\Rightarrow x+2y-3z=12$.

Example

$P, P_0 \in P$, $P_0(-1,2,3)$. Find P parallel to $P_0 = \{(x, y, z), 2x + 3z = 1\}$.

- Since $\overrightarrow{n_0} \perp P_0$, $\overrightarrow{n_0} = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}$.
- $\overrightarrow{n_0} \perp P$. So $P = 2x + 3z = 11$.

Lecture #13 – Tuesday, February 24, 2004

Example

Let U be a subspace of $\mathbf{R}^n$. Show that $\left(U^\perp\right)^\perp = U$.

- $\left(U^\perp\right)^\perp = \left\{X \in \mathbf{R}, X \cdot Y = 0 \forall Y \in U^\perp\right\}$.
- Let $Y \in \left(U^\perp\right)^\perp$. We want $Y \in U$. $\forall X_0 \in U^\perp$, $Y \cdot X_0 = 0 \Rightarrow Y \in U^\perp$.
- Also, since $\dim\left(\left(U^\perp\right)^\perp\right) + \dim\left(U^\perp\right) = n \Rightarrow \dim\left(\left(U^\perp\right)^\perp\right) = n - \dim\left(U^\perp\right)$ and $\dim(U) + \dim\left(U^\perp\right) = n \Rightarrow \dim(U) = n - \dim\left(U^\perp\right)$. This means $U^\perp \subset U$, and since $\dim(U) = \dim\left(\left(U^\perp\right)^\perp\right)$, $U = \left(U^\perp\right)^\perp$.

Example

Find an $n \times n$ matrix such that $U = \text{null}\, A = \left\{X \in \mathbf{R}^n, AX = 0\right\}$.

- Let $\{E_1, E_2, \cdots, E_k\}$ be an orthogonal basis of U. Let $\{F_1, F_2, \cdots, F_{n-k}\}$ be an orthogonal basis of $U^\perp$.
- Want: A such that $U = \left\{X \in \mathbf{R}^n, AX = 0\right\}$.
- Write $A = \begin{pmatrix} R_1 \\ \vdots \\ R_n \end{pmatrix}, R_j \in \mathbf{R}^n$. Now, $A = \begin{pmatrix} L_1^T \\ \vdots \\ L_n^T \end{pmatrix}, L_j$ are $1 \times n$ matrices.
- Want: A such that $AE_j = 0, j = 1, \cdots, k$. So $\begin{pmatrix} L_1^T \\ \vdots \\ L_n^T \end{pmatrix} E_j = 0 \Rightarrow \begin{pmatrix} L_1^T \cdot E_j \\ \vdots \\ L_n^T \cdot E_j \end{pmatrix} = 0 \Rightarrow L_k E_j = 0$. Since $L_k \in U^\perp$, so take $A = \begin{pmatrix} F_1^T \\ \vdots \\ F_{n-k}^T \\ 0 \\ \vdots \\ 0 \end{pmatrix}, F_i \in U^\perp$.

Example

Prove that $X \in \text{null}\, A$.

- Let $V = \text{null}\, A$.
- Want: $U \subset V$. Since $X \in U$, $AX = \begin{pmatrix} F_1{}^T \\ \vdots \\ F_{n-k}{}^T \\ 0 \\ \vdots \\ 0 \end{pmatrix} X = \begin{pmatrix} F_1{}^T X \\ \vdots \\ F_{n-k}{}^T X \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0$. So, $X \in \text{null}\, A = V$.
- Want: $V \subset U$. Since $X \in V$, $AX = \begin{pmatrix} F_1 \cdot X \\ \vdots \\ F_{n-k} \cdot X \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0$. So, $X \in \left(U^{\perp}\right)^{\perp} = U$.

Proposition

Let $U \subset V$ be subspaces in $\mathbf{R}^n$. Then $V^{\perp} \subset U^{\perp}$.

Proof:

- Let $X \in V^{\perp}$. So, $\forall Y \in V$, $X \cdot Y = 0$.
- Let $Z \subset U$. Then, $Z \subset V$. So, $X \cdot Z = 0 \Rightarrow X \in U^{\perp}$.

BEST APPROXIMATION

Theorem

Let A be an $m \times n$ matrix. Let $B \in \mathbf{R}^n$. We consider the system $AX = B, X \in \mathbf{R}^n$:

1) Any solution Z to the <u>normal equation</u> $A^T AZ = A^T B$ is a <u>best approximation</u> to a solution of $AX = B$, that is $\|B - AZ\| \le \|B - AX\|, X \in \mathbf{R}^n$.
2) If the columns of A are independent, then $A^T A$ is invertible, and Z is unique: $Z = \left(A^T A\right)^{-1} A^T B$.

Proof:

- Let $U = \left\{AX, X \in \mathbf{R}^n\right\} = \text{im}\, A = \text{col}\, A$. Then projection gives $AZ = P_U(B)$.
- Since $P_U(B) - B \in U^{\perp}$, so $\forall Y \in U$, $\left(P_U(B) - B\right) \cdot Y = 0 \Rightarrow \left(P_U(B) - B\right) \cdot (AX) = 0$.
- $\left(P_U(B) - B\right) \cdot (AX) = (AX)^T \left(P_U(B) - B\right) = X^T \left[A^T \left(P_U(B) - B\right)\right] = X \cdot \left[A^T \left(P_U(B) - B\right)\right] = 0$.
- Remark: If $X \cdot Y = 0$ for any X, then $Y = 0$. So, $A^T \left(P_U(B) - B\right) = 0 \Rightarrow A^T (AZ - B) = 0 \Rightarrow A^T AZ = A^T B$.

- Remark: If A is invertible, we know that the solution of $AX = B$ is $X = A^{-1} B$.
- The solution of the normal equation is $A^T AZ = A^T B$. Since A^T is invertible if A is invertible, $\left(A^T\right)^{-1} A^T AZ = \left(A^T\right)^{-1} A^T B \Rightarrow AZ = B \Rightarrow Z = A^{-1} B$.

<u>Remark</u>: If Z is the best approximation, $A^T AZ = A^T B$. If $Z_0 \in \text{null} A$, that is $AZ_0 = 0$, then $Z + Z_0$ is also a best approximation. $A^T A(Z + Z_0) = A^T(AZ + AZ_0) = A^T(AZ) = A^T AZ = A^T B$.

Example

Find the best approximation for $\begin{cases} x + y = 5 \\ 2x - y = 1 \\ -x = 0 \end{cases}$.

- Let $AX = B$, $A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 0 \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$, $B = \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix}$.
- Consider $A^T AZ = A^T B$. $A^T A = \begin{pmatrix} 6 & -1 \\ -1 & 2 \end{pmatrix} \Rightarrow (A^T A)^{-1} = \frac{1}{11}\begin{pmatrix} 2 & 1 \\ 1 & 6 \end{pmatrix}$. So,

  $Z = (A^T A)^{-1} A^T B = \begin{pmatrix} \frac{18}{11} \\ \frac{31}{11} \end{pmatrix} \Rightarrow \begin{cases} x = \frac{18}{11} \\ y = \frac{31}{11} \end{cases}$.

Example

Find U such that $U = \text{im} A = \{AX, X \in \mathbf{R}^2\}$.

- Find an orthogonal basis of U:
  - $U = \text{col} A = \text{span}\left\{\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}\right\}$.
  - Apply the Gram-Schmidt Orthogonalisation:
    - $E_1 = X_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$.
    - $E_2 = X_2 - \frac{X_2 \cdot E_1}{\|E_1\|^2} E_1 = \begin{pmatrix} \frac{3}{2} \\ \frac{3}{2} \\ -1 \end{pmatrix}$.
  - So, $\{E_1, E_2\}$ is an orthogonal basis of U.
- $P_U(B) = \frac{B \cdot E_1}{\|E_1\|^2} E_1 + \frac{B \cdot E_2}{\|E_2\|^2} E_2 = \begin{pmatrix} \frac{49}{11} \\ \frac{8}{11} \\ -\frac{18}{11} \end{pmatrix}$.
- Z, the best approximation is such that $AZ = P_U(B)$.

- Let $Z = \begin{pmatrix} x \\ y \end{pmatrix}$. We have to solve $AZ = P_U(B)$. $\begin{cases} x + y = \frac{49}{11} \\ 2x - y = \frac{5}{11} \\ -x = -\frac{18}{11} \end{cases} \Rightarrow \begin{cases} x = \frac{18}{11} \\ y = \frac{31}{11} \end{cases}$.

LEAST SQUARE APPROXIMATION

Definition

Let $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ be the given data. Let $m \in \mathbf{N}$. We look for a polynomial $p(x)$ of degree *m* such that $(y_1 - p(x_1))^2 + (y_2 - p(x_2))^2 + \cdots + (y_n - p(x_n))^2$ is minimum. The $p(x)$ satisfying this condition is called the <u>least square approximating polynomial</u> of degree *m*.

- Let $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, $p(x) = \begin{pmatrix} p(x_1) \\ \vdots \\ p(x_n) \end{pmatrix}$.
- $(y_1 - p(x_1))^2 + (y_2 - p(x_2))^2 + \cdots + (y_n - p(x_n))^2 = \|Y - p(x)\|^2$. In particular, we want to minimize $\|Y - p(x)\|$.
- Write $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$, $R = \begin{pmatrix} a_0 \\ \vdots \\ a_m \end{pmatrix} \in \mathbf{R}^{m+1}$, $M_{n \times (m+1)} = \begin{pmatrix} 1 & x_1 & x_1^{\,2} & \cdots & x_1^{\,m} \\ 1 & x_2 & x_2^{\,2} & \cdots & x_2^{\,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^{\,2} & \cdots & x_n^{\,m} \end{pmatrix}$. So now, $p(x) = MR$.
- We look for $R \in \mathbf{R}^{m+1}$ that minimizes $\|Y - MR\|$.

Theorem

1) If $Z = \begin{pmatrix} z_0 \\ \vdots \\ z_n \end{pmatrix}$ is a solution to the normal equation $M^T M Z = M^T Y$, then $p(x) = z_0 + z_1 x + z_2 x^2 + \cdots + z_m x^m$ is a least square approximation polynomial of degree *m*.
2) If at least $m+1$ of the x_j are distinct, then $M^T M$ is invertible, so $Z = (M^T M)^{-1} M^T Y$ is unique.

Proof (#1): This is immediate and is given by the theorem above.

Proof (#2):

- Suppose $x_1, \cdots, x_{m+1}$ are distinct. We want to prove that the columns of *M* are linearly independent.

- $$t_0\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}+t_1\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}+t_2\begin{pmatrix}x_1{}^2\\ \vdots\\ x_n{}^2\end{pmatrix}+\cdots+t_m\begin{pmatrix}x_1{}^m\\ \vdots\\ x_n{}^m\end{pmatrix}=0\Rightarrow\begin{cases}t_0+t_1x_1+t_2x_1{}^2+\cdots+t_mx_1{}^m=0\\ t_0+t_1x_2+t_2x_2{}^2+\cdots+t_mx_2{}^m=0\\ \qquad\qquad\vdots\\ t_0+t_1x_n+t_2x_n{}^2+\cdots+t_mx_n{}^m=0\end{cases}.$$
- Let $q(x)=t_0+t_1x+t_2x^2+\cdots+t_mx^m=0$. So $q(x_j)=0,\ j=1,\cdots,n$.
- But $x_1,\cdots,x_{m+1}$ are distinct. So q has at least $m+1$ distinct roots. So since the degree of q is m, $q=0$. Therefore, $t_0=t_1=\cdots=t_m=0$.
- Therefore, the columns of M are linearly independent. So, M^TM is invertible.

Example

Suppose given $(-1,1)$, $(1,2)$, $(0,3)$. Find the least square approximating polynomial of degree 1.

- Want: $p(x)=a_0+a_1x$ such that $(1-p(-1))^2+(2-p(1))^2+(3-p(0))^2$ is minimum.
- Let $R=\begin{pmatrix}a_0\\ a_1\end{pmatrix}$, $M=\begin{pmatrix}1&-1\\ 1&1\\ 1&0\end{pmatrix}=\begin{pmatrix}1&x_1\\ 1&x_2\\ 1&x_3\end{pmatrix}$, $Y=\begin{pmatrix}1\\ 2\\ 3\end{pmatrix}$. Find R such that $\|MR-Y\|$ is minimum.
- Let $U=\operatorname{col}M=\operatorname{span}\left\{\begin{pmatrix}1\\ 1\\ 1\end{pmatrix},\begin{pmatrix}-1\\ 1\\ 0\end{pmatrix}\right\}$. Apply Gram-Schmidt Orthogonalisation to find an orthogonal basis:
  - $E_1=\begin{pmatrix}-1\\ 1\\ 0\end{pmatrix}$.
  - $E_2=X_2-\dfrac{X_2\cdot E_1}{\|E_1\|^2}E_1=\begin{pmatrix}1\\ 1\\ 1\end{pmatrix}$.
- $$P_U(Y)=\frac{Y\cdot E_1}{\|E_1\|^2}E_1+\frac{Y\cdot E_2}{\|E_2\|^2}E_2=\frac{1}{2}\begin{pmatrix}-1\\ 1\\ 0\end{pmatrix}+\frac{6}{14}\begin{pmatrix}1\\ 1\\ 1\end{pmatrix}=\begin{pmatrix}-\frac{1}{14}\\ \frac{13}{14}\\ \frac{3}{7}\end{pmatrix}.$$
- Now solve the system $$MR=P_U(Y)\Rightarrow\begin{cases}a_0-a_1=-\dfrac{1}{14}\\ a_0-a_1=\dfrac{13}{14}\\ a_0=\dfrac{3}{7}\end{cases}\Rightarrow\begin{cases}a_0=\dfrac{3}{7}\\ a_1=\dfrac{7}{14}=\dfrac{1}{2}\end{cases}.$$
- Therefore, the least square approximating polynomial is $p(x)=\dfrac{3}{7}+\dfrac{x}{2}$.

Lecture #14 – Thursday, February 26, 2004

Definition

Let $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ be the given data. Let $f_0(x), f_1(x), \cdots, f_n(x)$ be the given functions. We look for $f(x) = m_0 f_0(x) + m_1 f_1(x) + \cdots m_m f_m(x)$ such that
$(y_1 - f(x_1))^2 + (y_2 - f(x_2))^2 + \cdots + (y_n - f(x_n))^2$ is minimum. f is the least approximating function.

Example

Given data $(1,-2)$, $(2,0)$, $(0,3)$ and $f_0(x) = x$, $f_1(x) = 3^x$, find the least approximating function.

- Let $M = \begin{pmatrix} f_0(1) & f_1(1) \\ f_0(2) & f_1(2) \\ f_0(0) & f_1(0) \end{pmatrix} = 0\begin{pmatrix} 1 & 3 \\ 2 & 9 \\ 0 & 1 \end{pmatrix}$, $Y = \begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}$, $Z = \begin{pmatrix} z_0 \\ z_1 \end{pmatrix}$. Have to solve $M^T MZ = M^T Y$.
- $M^T M = \begin{pmatrix} 5 & 21 \\ 21 & 91 \end{pmatrix}$. So $M^T MZ = M^T Y \Rightarrow \begin{pmatrix} 5 & 21 \\ 21 & 91 \end{pmatrix}\begin{pmatrix} z_0 \\ z_1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 9 & 1 \end{pmatrix}\begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix} \Rightarrow \begin{pmatrix} 5z_0 + 21z_1 \\ 21z_0 + 91z_1 \end{pmatrix} = \begin{pmatrix} -2 \\ -3 \end{pmatrix}$.
- So, $\begin{cases} 5z_0 + 21z_1 = -2 \\ 21z_0 + 91z_1 = -3 \end{cases} \Rightarrow \begin{cases} z_0 = -\dfrac{17}{2} \\ z_1 = \dfrac{21}{40} \end{cases}$.
- Therefore, the least approximating function is $f(x) = -\dfrac{17}{2}x + \dfrac{27}{40}3^x$.

Subspaces Associated To Matrices and Orthogonality

Let A be an $m \times n$ matrix, and $A = \begin{pmatrix} C_1 & \cdots & C_n \end{pmatrix}$ or $A = \begin{pmatrix} R_1 \\ \vdots \\ R_m \end{pmatrix}$. The subspaces associated to A are:

- $\text{null}(A) = \{X \in \mathbf{R}^n, AX = 0\}$.
- $\text{col}(A) = \text{im}(A) = \text{span}(C_j) = \{Y \in \mathbf{R}^n, Y = AX\}$.
- $\text{row}(A) = \text{span}(R_j)$.

Proposition

1) $\text{null}(A) = (\text{row}(\text{A}))^{\perp}$ – the rows are considered as vectors.
2) $\text{null}(A^T) = (\text{col}(\text{A}))^{\perp}$.

Proof 1:

- Dimension:
    - Let $r = \text{rank}(A) = \dim(\text{col}(A)) = \dim(\text{row}(A))$.
    - $r + \dim(\text{null}(A)) = m \Rightarrow \dim(\text{null}(A)) = m - r$.
    - $\dim(\text{row}(A)) + \dim((\text{row}(A))^{\perp}) = m \Rightarrow \dim((\text{row}(A))^{\perp}) = m - r$.

- So, $\dim\left((\mathrm{row}(A))^{\perp}\right)=\dim(\mathrm{null}(A))$.
- $\mathrm{null}(A)\subset(\mathrm{row}(A))^{\perp}$:
  - Let $AX\in\mathrm{null}(A)$. So $AX=0$.
  - Let $R_1,\cdots,R_m$ be the rows of A. Then $AX=\begin{pmatrix}R_1\\ \vdots\\ R_m\end{pmatrix}X=\begin{pmatrix}R_1X\\ \vdots\\ R_mX\end{pmatrix}=0\Rightarrow R_jX=0$.
  - Let $R_j{}^T=\begin{pmatrix}r_{j1} & \cdots & r_{jn}\end{pmatrix}^T=Y_j=\begin{pmatrix}m_{j1}\\ \vdots\\ m_{jn}\end{pmatrix}$.
  - Now, $AX=\begin{pmatrix}Y_1{}^TX\\ \vdots\\ Y_m{}^TX\end{pmatrix}=\begin{pmatrix}Y_1\cdot X\\ \vdots\\ Y_m\cdot X\end{pmatrix}=0$. So X is orthogonal to Y_j, so X is orthogonal to $\mathrm{span}(Y_j)=\mathrm{row}(A)$. Therefore $\mathrm{null}(A)\subset(\mathrm{row}(A))^{\perp}$.
- Since $\dim\left((\mathrm{row}(A))^{\perp}\right)=\dim(\mathrm{null}(A))$ and $\mathrm{null}(A)\subset(\mathrm{row}(A))^{\perp}$, $\mathrm{null}(A)=(\mathrm{row}(\mathrm{A}))^{\perp}$.

Proof 2:

- We have $\mathrm{null}(A)=(\mathrm{row}(\mathrm{A}))^{\perp}\Rightarrow\mathrm{null}\left(A^T\right)=\left(\mathrm{row}\left(\mathrm{A}^T\right)\right)^{\perp}\Rightarrow\mathrm{null}\left(A^T\right)=(\mathrm{col}(\mathrm{A}))^{\perp}$.

Example

Let $A=\begin{pmatrix}1 & -2 & 1 & 0 & 2\\ 1 & -1 & 4 & 1 & 3\\ -1 & 3 & 2 & 1 & -1\\ 2 & -3 & 5 & 1 & 5\end{pmatrix}$. Calculate $\mathrm{null}(A)$ and $\mathrm{row}(A)$, and prove that $\mathrm{null}(A)$ is orthogonal to $\mathrm{row}(A)$.

- $\mathrm{null}(A)=\{X, AX=0\}$. Let $X=\begin{pmatrix}x_1\\ x_2\\ x_3\\ x_4\\ x_5\end{pmatrix}$. $AX=0\Leftrightarrow\begin{cases}x_1-x_2+x_3+2x_5=0\\ x_1-x_2+4x_3+x_4+3x_5=0\\ -x_1+3x_2+2x_3+x_4-x_5=0\\ 2x_1-3x_2+5x_3+x_4+5x_5=0\end{cases}$. So,

  $X=s\begin{pmatrix}-7\\ -3\\ 1\\ 0\\ 0\end{pmatrix}+t\begin{pmatrix}-2\\ -1\\ 0\\ 1\\ 0\end{pmatrix}+r\begin{pmatrix}-1\\ -1\\ 0\\ 0\\ 1\end{pmatrix}$.
- By row reduction, $A\to B=\begin{pmatrix}1 & 0 & 7 & 2 & 4\\ 0 & 1 & 3 & 1 & 1\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}$. So now, $\mathrm{row}(A)=\mathrm{span}\left\{\begin{pmatrix}1\\ 0\\ 7\\ 2\\ 4\end{pmatrix},\begin{pmatrix}0\\ 1\\ 3\\ 1\\ 1\end{pmatrix}\right\}$.

- $\text{null}(A)$ is orthogonal to $\text{row}(A)$:
  - $\dim(\text{row}(A)) = 2 \Rightarrow \dim\left((\text{row}(A))^{\perp}\right) = 5-2 = 3$. Also, $\dim(\text{null}(A))$.
  - For every vector in $\text{null}(A)$ and $\text{row}(A)$, the dot product is 0. So, $\text{null}(A)$ is orthogonal to $\text{row}(A)$.

Example

Find $a \in \mathbf{R}$ such that $\begin{pmatrix} a^2 \\ -3a \\ -2 \end{pmatrix} \in \text{span}\left\{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}\right\}$.

- Find an orthogonal basis:
  - $E_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.
  - $E_2 = X_2 - \dfrac{X_2 \cdot E_1}{\|E_1\|^2} E_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \dfrac{5}{2}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.
  - $E_3 = X_3 - \dfrac{X_3 \cdot E_1}{\|E_1\|^2} E_1 - \dfrac{X_3 \cdot E_2}{\|E_2\|^2} E_2 = \begin{pmatrix} 1 \\ 5 \\ 4 \end{pmatrix} - \dfrac{9}{2}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \dfrac{16}{11}\begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.

Lecture #15 – Tuesday, March 2, 2004

TRANSFORMATIONS

Definitions

1) A <u>transformation *T*</u> from $\mathbf{R}^n$ to $\mathbf{R}^m$ is a rule $T(X)$ that assigns to every $X \in \mathbf{R}^n$ a uniquely determined vector in $\mathbf{R}^m$. We write $T : \mathbf{R}^n \to \mathbf{R}^m$ or $\mathbf{R}^n \xrightarrow{T} \mathbf{R}^m$.
2) If $T : \mathbf{R}^n \to \mathbf{R}^m$, $\mathbf{R}^n$ is the <u>domain of *T*</u>, and $\mathbf{R}^m$ is the <u>co-domain of *T*</u>.

Examples

1) Find the domain and co-domain of $T_1 : \mathbf{R}^2 \to \mathbf{R}^3$ such that $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x^2 \\ -y \\ 2x+y \end{pmatrix}$.
   - Domain of T_1: $\mathbf{R}^2$.
   - Co-domain of T_1: $\mathbf{R}^3$.
2) Describe $T_2 : \mathbf{R}^2 \to \mathbf{R}^2$, a reflection in the *y*-axis.
   - $T_2(x, y) = (-x, y)$.

3) Describe $T_3 : \mathbf{R}^3 \to \mathbf{R}^3$, the projection on the (x, y)-plane.
   - $T_3(x, y, z) = (x, y, 0)$.

Example

Let A be an $m \times n$ matrix. $T_A : \mathbf{R}^n \to \mathbf{R}^m$ such that $X \to AX$.

- T_A is the transformation induced by A.
- If $A = 0$, $T_A(X) = 0$ for any $X \in \mathbf{R}^n$.
- If $A = I_n$, $T_A(X) = I_n X = X$. $T_A(X) = 1_{\mathbf{R}^n}$.

LINEAR TRANSFORMATIONS

Definition

Let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a transformation. T is linear iff for any vector $X, Y \in \mathbf{R}^n$ and scalar $a \in \mathbf{R}$ we have:
   - (T1): $T(X+Y) = T(X) + T(Y)$ – "T preserves the addition".
   - (T2): $T(aX) = aT(X)$ – "T preserves the multiplication by a scalar".

- Remark: For $T : \mathbf{R}^n \to \mathbf{R}^m$, T is linear if $T(0) = 0$ because $a = 0$ in (T2).

Example

Let $T_Z : \mathbf{R}^n \to \mathbf{R}^n$ such that $X \to X + Z$. Is T_Z linear?

- If T_Z is linear, $T(0) = 0$. Since $T_Z(0) = 0 + Z$, so if T_Z is linear, $Z = 0$.

Example

Consider $T_1 : \mathbf{R}^2 \to \mathbf{R}^3$ such that $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 2x+y \\ x-y \\ x \end{pmatrix}$. Is T linear?

- Condition (T1):
   - Let, $Y = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$.
   - $T_1(X+Y) = T_1\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} = \begin{pmatrix} 2(x_1 + x_2) + (y_1 + y_2) \\ (x_1 + x_2) - (y_1 + y_2) \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + y_1 + 2x_2 + y_2 \\ x_1 - y_1 + x_2 - y_2 \\ x_1 + x_2 \end{pmatrix}$.
   - $T_1(X) + T(Y) = T_1\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + T_1\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + y_1 \\ x_1 - y_1 \\ x_1 \end{pmatrix} + \begin{pmatrix} 2x_2 + y_2 \\ x_2 - y_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + y_1 + 2x_2 + y_2 \\ x_1 - y_1 + x_2 - y_2 \\ x_1 + x_2 \end{pmatrix}$.
   - So $T_1(X+Y) = T_1(X) + T_1(Y)$ and (T1) is satisfied.
- Condition (T2):

- Let $X=\begin{pmatrix}x\\y\end{pmatrix}$, $a\in\mathbf{R}$.
- $T_1(aX)=T_1\begin{pmatrix}ax\\ay\end{pmatrix}=\begin{pmatrix}2ax+ay\\ax-ay\\ax\end{pmatrix}=a\begin{pmatrix}2x+y\\x-y\\x\end{pmatrix}=aT_1(X)$.
- So $T_1(aX)=aT_1(X)$ and (T2) is satisfied.

- Since (T1) and (T2) are satisfied, T_1 is linear.

Examples

1) Is $T_2:\mathbf{R}^2\to\mathbf{R}^2$ such that $\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}x+1\\y\end{pmatrix}$ linear?
   - $T_2(0)=\begin{pmatrix}1\\0\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix}$. So T_2 is not linear.

2) Is $T_3:\mathbf{R}^2\to\mathbf{R}$ such that $\begin{pmatrix}x\\y\end{pmatrix}\to\left(x^2+y\right)$ linear?
   - Check (T2): Let $X=\begin{pmatrix}x\\y\end{pmatrix}$, $a\in\mathbf{R}$. $T_3(aX)=T_3\begin{pmatrix}ax\\ay\end{pmatrix}=a^2x^2+ay$, and $aT_3(X)=a\left(x^2+y\right)=ax^2+ay$. Pick $a=2$. $T_3(2X)=4x^2+2y$, but $2T_3(X)=2x^2+2y$.
   - So, since (T2) is not satisfied, T_3 is not linear.

LINEAR COMBINATIONS AND LINEAR TRANSFORMATIONS

Theorem

If $T:\mathbf{R}^n\to\mathbf{R}^m$ is a linear transformation. Then $T(a_1X_1+\cdots+a_kX_k)=a_1T(X_1)+\cdots+a_kT(X_k)$ for any $a_j\in\mathbf{R}$ and $X_j\in\mathbf{R}^n$.

Proof: We use induction on *k*.

- Base case: $k=1$.
  - $T(a_1X_1)=a_1T(X_1)$ since *T* is linear.
- Assume true for *k*. Prove for $k+1$.
  - $T(a_1X_1+\cdots+a_kX_k+a_{k+1}X_{k+1})=T(a_1X_1+\cdots+a_kX_k)+T(a_{k+1}X_{k+1})$ since *T* is linear.
  - $T(a_1X_1+\cdots+a_kX_k)+T(a_{k+1}X_{k+1})=a_1T(X_1)+\cdots+a_kT(X_k)+T(a_{k+1}X_{k+1})$ by induction hypothesis.
  - $a_1T(X_1)+\cdots+a_kT(X_k)+T(a_{k+1}X_{k+1})=a_1T(X_1)+\cdots+a_kT(X_k)+a_{k+1}T(X_{k+1})$ since *T* is linear.
  - So, $T(a_1X_1+\cdots+a_kX_k+a_{k+1}X_{k+1})=a_1T(X_1)+\cdots+a_kT(X_k)+a_{k+1}T(X_{k+1})$.
- So, the theorem is proven.

Example

Let *T* linear be $T:\mathbf{R}^3\to\mathbf{R}^2$ such that $T\begin{pmatrix}1\\2\end{pmatrix}=\begin{pmatrix}2\\-1\end{pmatrix}$ and $T\begin{pmatrix}3\\-1\end{pmatrix}=\begin{pmatrix}1\\-4\end{pmatrix}$. Compute $T\begin{pmatrix}2\\-3\end{pmatrix}$.

- If $\begin{pmatrix} 2 \\ -3 \end{pmatrix} = a\begin{pmatrix} 1 \\ 2 \end{pmatrix} + b\begin{pmatrix} 3 \\ -1 \end{pmatrix}$, then $T\begin{pmatrix} 2 \\ -3 \end{pmatrix} = aT\begin{pmatrix} 1 \\ 2 \end{pmatrix} + bT\begin{pmatrix} 3 \\ -1 \end{pmatrix}$.
- Solving for a and b, we obtain $a = -1$, $b = 1$.
- So, $T\begin{pmatrix} 2 \\ -3 \end{pmatrix} = -T\begin{pmatrix} 1 \\ 2 \end{pmatrix} + T\begin{pmatrix} 3 \\ -1 \end{pmatrix} = -\begin{pmatrix} 2 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -4 \end{pmatrix} = \begin{pmatrix} -1 \\ -3 \end{pmatrix}$.

Theorem

Let $\mathbf{R}^n = \operatorname{span}\{X_1, \cdots, X_k\}, k \ge n$. Let S and T be linear transformations $\mathbf{R}^n \to \mathbf{R}^m$. Then:

1) T is defined by $T(X_j), j = 1, \cdots, k$.
2) $T = S \Leftrightarrow T(X_j) = S(X_j), j = 1, \cdots, k$.

Proof of 1:

- Let $X \in \mathbf{R}^n$. Let T be linear such that $T(X_j)$ are known.
- Since $X = a_1 X_1 + \cdots + a_k X_k$, so $T(X) = T(a_1 X_1 + \cdots + a_k X_k) = a_1 T(X_1) + \cdots + a_k T(X_k)$ is known.

Proof of 2:

- Assume $T = S$, $T(X) = S(X)$ for any $X \in \mathbf{R}^n$. So for $X = X_j$, $T(X_j) = S(X_j)$.
- Assume $T(X_j) = S(X_j), j = 1, \cdots, k$. Let $X \in \mathbf{R}^n$, so $X = a_1 X_1 + \cdots + a_k X_k$. So $T(X) = a_1 T(X_1) + \cdots + a_k T(X_k) = a_1 S(X_1) + \cdots + a_k S(X_k) = S(a_1 X_1 + \cdots + a_k X_k) = S(X)$.

- <u>Remark</u>: T is defined by $T(E_j)$, $\{E_1, \cdots, E_n\}$ is a basis of the domain of T.

Example

Let $V, W \in \mathbf{R}^3$, $V = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$, $W = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$. Define $V \,\mathrm{n}\, W = \begin{pmatrix} y_1 z_2 - y_2 z_1 \\ z_1 x_2 - z_2 x_1 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$. Show that for U, V, W in $\mathbf{R}^3$, $U \,\mathrm{n}\, (V \,\mathrm{n}\, W) = (U \cdot W)V - (U \cdot V)W$.

- Let $T(U) = U \,\mathrm{n}\, (V \,\mathrm{n}\, W)$ and $S(U) = (U \cdot W)V - (U \cdot V)W$ for V, W fixed. It can be proven that T, S are linear.
- Using the theorem above, we only need to prove that $T(E_j) = S(E_j)$ where $\{E_1, E_2, E_3\}$ is the standard basis of $\mathbf{R}^3$.
    - $S(E_1) = (E_1 \cdot W)V - (E_1 \cdot V)W = x_2 \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} - x_1 \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} 0 \\ x_2 y_1 - x_1 y_2 \\ x_2 z_1 - x_1 z_2 \end{pmatrix}$.
    - $T(E_1) = E_1 \,\mathrm{n}\, (V \,\mathrm{n}\, W) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mathrm{n} \begin{pmatrix} y_1 z_2 - y_2 z_1 \\ z_1 x_2 - z_2 x_1 \\ x_1 y_2 - x_2 y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ x_2 y_1 - x_1 y_2 \\ x_2 z_1 - x_1 z_2 \end{pmatrix}$.
    - So $T(E_1) = S(E_1)$. The same goes for E_2 and E_3, so $T = S$.
- Therefore, $U \,\mathrm{n}\, (V \,\mathrm{n}\, W) = (U \cdot W)V - (U \cdot V)W$.

MATRIX OF A LINEAR TRANSFORMATION

Proposition

Let A be an $m\times n$ matrix. $T_A : \mathbf{R}^n \to \mathbf{R}^m, X \to AX$ is linear. (Proof: Properties of matrix multiplication).

Theorem

Let $T_A : \mathbf{R}^n \to \mathbf{R}^m$ be a transformation. Then:

1) T is linear iff it is a matrix transformation.
2) If T is linear, then T is induced by the (unique) matrix A defined by (in terms of its columns) $A = [T(E_1) \quad \cdots \quad T(E_n)]$, where $\{E_1, \cdots, E_n\}$ is the standard basis of $\mathbf{R}^n$.

Proof of 1:

- If $T = T_A$ then T is linear by proposition.
- Let T be a linear transformation. Let $A = [T(E_1) \quad \cdots \quad T(E_n)]$ an $n\times n$ matrix. We want to prove $T = T_A$.
    - By theorem above, it is enough to prove that $T(E_j) = T_A(E_j)$, $j = 1, \cdots, n$.
    - $E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. $T_A(E_1) = [T(E_1) \quad \cdots \quad T(E_n)] \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = T(E_1)$.
    - Continuing this way, we obtain that $T(E_j) = T_A(E_j)$, $j = 1, \cdots, n$, so $T = T_A$.

Proof of 2:

- Assume $T = T_A = T_B$.
- $T_A(X) = T_B(X) \forall X \in \mathbf{R}^n$. So $AX = BX$.
- $B = BI_n = B(E_1 \quad \cdots \quad E_n) = (BE_1 \quad \cdots \quad BE_n) = (AE_1 \quad \cdots \quad AE_n) = AI_n = A$.
- So A is unique.

Corollary

If $AX = BX$ for any $X \in \mathbf{R}^n$, then $A = B$.

Examples

1) Let $T_1 : \mathbf{R}^3 \to \mathbf{R}^3, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} x+y \\ z-x+2y \\ 3x-z \end{pmatrix}$. T_1 is linear. Find matrix A_1 of T_1.

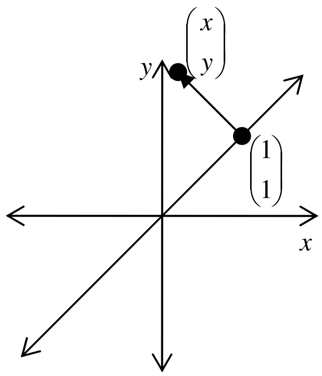
    - $A_1 = [T(E_1) \quad T(E_2) \quad T(E_3)] = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ 3 & 0 & -1 \end{pmatrix}$.

2) Let $T_2 : \mathbf{R}^2 \to \mathbf{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} -x \\ y \end{pmatrix}$ a reflection in y-axis. T_2 is linear. Find matrix A_2 of T_2.

    - $A_2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

Example

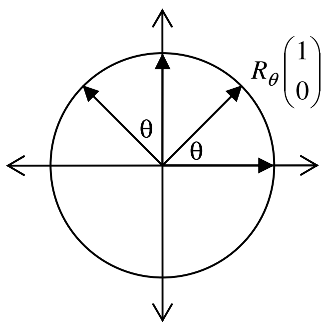
Let $P:\mathbf{R}^2 \to \mathbf{R}^2$ the orthogonal projection on $\operatorname{span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$. P is linear. Find matrix M_P of P.



- $P(X)=\dfrac{X\cdot X_0}{\|X_0\|}X_0 \Rightarrow P\begin{pmatrix}x\\y\end{pmatrix}=\dfrac{\begin{pmatrix}x\\y\end{pmatrix}\cdot\begin{pmatrix}1\\1\end{pmatrix}}{2}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}\frac{x+y}{2}\\\frac{x+y}{2}\end{pmatrix}$.
- So $P:\mathbf{R}^2\to\mathbf{R}^2$ such that $P\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}\frac{x+y}{2}\\\frac{x+y}{2}\end{pmatrix}$.
- So the matrix of P is $M_P=\left[P\begin{pmatrix}1\\0\end{pmatrix}\quad P\begin{pmatrix}0\\1\end{pmatrix}\right]=\begin{pmatrix}\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}\end{pmatrix}$.

Example

Let $R_\theta:\mathbf{R}^2\to\mathbf{R}^2$ a rotation of centre 0 and of angle θ (ccw). R_θ is linear. Find matrix M_θ of R_θ.



- $R_\theta\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix},\ R_\theta\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}-\sin\theta\\\cos\theta\end{pmatrix}$.
- So M_θ the matrix of R_θ is given by $M_\theta=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}$.

Lecture #16 – Thursday, March 4, 2004

Composition of Linear Transformation

Definition

Let S, T be two transformation such that $\operatorname{dom}(S)=\operatorname{codom}(T)$ and $\mathbf{R}^n\xrightarrow{T}\mathbf{R}^m\xrightarrow{S}\mathbf{R}^p$. We define <u>composite of S and T</u> $(S\circ T)(X)=S(T(X))$ for $X\in\mathbf{R}^n$. $S\circ T:\mathbf{R}^n\to\mathbf{R}^p$.

Theorem

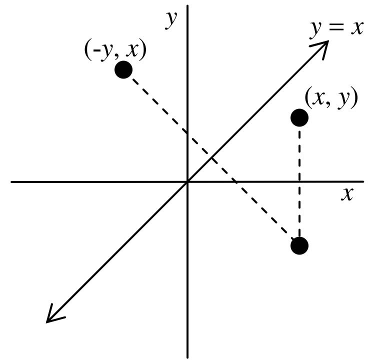
Let $T:\mathbf{R}^n\to\mathbf{R}^m$ and $S:\mathbf{R}^m\to\mathbf{R}^p$ be linear transformations with matrix B and A respectively. Then $S\circ T:\mathbf{R}^n\to\mathbf{R}^p$ is linear and its matrix is AB.

<u>Proof</u>: We have $(S\circ T)(X)=S(T(X))=S(BX)=A(BX)=(AB)X$.

Example

Let T be the reflection in x-axis and S be the reflection with respect to the line $y = x$. What is the matrix for $S \circ T$?

- Matrix of T: $B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.
- Matrix of S: $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.
- $S \circ T$ has matrix $AB = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.



Inverse of Linear Transformation

Definition

Let $T : \mathbf{R}^n \to \mathbf{R}^n$ be a linear transformation. Then T is invertible iff $\exists S : \mathbf{R}^n \to \mathbf{R}^n$ such that $S \circ T = 1_{\mathbf{R}^n}$ and $T \circ S = 1_{\mathbf{R}^n}$.

Theorem

Let $T : \mathbf{R}^n \to \mathbf{R}^n$ be linear with matrix A. Then T is invertible iff A is invertible. In this case, the matrix of T^{-1} is A^{-1}, and so T is unique.

<u>Proof</u>:

- If A is invertible, $T_A \circ T_{A^{-1}} = (AA^{-1})X = X$ and $T_{A^{-1}} \circ T_A = (A^{-1}A)X = X$, so T is invertible.
- If T is invertible, let T' be an inverse (T' is linear because for $\tilde{X} = T'(X)$ and $\tilde{Y} = T'(Y)$, $T'(X+Y) = T'(T(\tilde{X}) + T(\tilde{Y})) = T'(T(\tilde{X} + \tilde{Y})) = T' \circ T(\tilde{X} + \tilde{Y}) = \tilde{X} + \tilde{Y} = T'(X) + T'(Y)$, and $T'(aX) = T'(aT(\tilde{X})) = T'(T(a\tilde{X})) = (T' \circ T)(a\tilde{X}) = a\tilde{X} = aT'(X)$). let B be the matrix of T'. then $T \circ T' = 1_{\mathbf{R}^n} \Rightarrow AB = I_n$ and $T' \circ T = 1_{\mathbf{R}^n} \Rightarrow BA = I_n$. So $B = A^{-1}$ and A is invertible, and thus T' is unique.

Lecture #17 – Tuesday, March 9, 2004

The Determinant of 1×1 and 2×2 Matrices

Definition

Let $A = [a], a \in \mathbf{R}$ be an 1×1 matrix. Then $\det A = a$.

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, a, b, c, d \in \mathbf{R}$ be an 2×2 matrix. Then $\det A = |A| = ad - bc$.

Examples

1) The determinant of $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ is $\det A = (1)(4) - (2)(3) = -2$.
2) The determinant of $B = \begin{pmatrix} -1 & 5 \\ 3 & -2 \end{pmatrix}$ is $\det B = (-1)(-2) - (3)(5) = -13$.

Proposition

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then A is invertible iff $\det A \neq 0$. Moreover, in this case $A^{-1} = \dfrac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Proof:

- Prove: $\det A \neq 0 \Rightarrow A$ is invertible.
  - Define $B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Then $AB = (\det A) I_2 \Leftrightarrow A\left(\dfrac{1}{\det A} B\right) = I_2 \Rightarrow A$ is invertible and $A^{-1} = \dfrac{1}{\det A} B$.
- Prove: A is invertible $\Rightarrow \det A \neq 0$.
  - Assume $\det A = 0$. Want A cannot be invertible.
  - Define $B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Then $AB = (\det A) I_2 = 0$.
  - But $A^{-1}(AB) = B = 0 \Rightarrow a = b = c = d = 0 \Rightarrow A = 0$. Contradiction!

Example

Compute $\det A$ and A^{-1} when it exists.

1) $A_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.
   - $\det(A_1) = (1)(4) - (3)(2) = -2$, so A_1 is invertible.
   - $A_1^{-1} = \dfrac{1}{\det A}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$.

2) $A_2 = \begin{pmatrix} -\frac{3}{2} & 3 \\ 2 & -4 \end{pmatrix}$.
   - $\det(A_2) = \left(-\dfrac{3}{2}\right)(-4) - (3)(2) = 0$, so A_2 is not invertible.

Example

Let $A = \begin{pmatrix} 1 & a \\ -2 & 3 \end{pmatrix}$. For which value of $a \in \mathbf{R}$ is the matrix A invertible? Compute A^{-1} for those values.

- $\det A = (1)(3) - (-2)(a) = 3 + 2a$. So $\det A \neq 0$ iff $a \neq -\dfrac{3}{2}$.
- If $a \neq -\dfrac{3}{2}$, $A^{-1} = \dfrac{1}{3 - 2a}\begin{pmatrix} 3 & -a \\ 2 & 1 \end{pmatrix}$.

THE DETERMINANT FOR $N\times N$ MATRICE

Definition

Let A be $n\times n$ matrix. Let A_{ij} be the $(n-1)\times(n-1)$ matrix formed from A by deleting row i and column j.
Then the (i, j)-<u>cofactor</u> is $C_{ij}(A)=(-1)^{i+j}\det A_{ij}\in \mathbf{R}$.

- <u>Remark</u>: If $i+j$ is even, $(-1)^{i+j}=1$. If $i+j$ is odd, $(-1)^{i+j}=-1$.

Example

Let $A=\begin{pmatrix}1 & -2 & 3\\ 4 & 1 & -1\\ 0 & 2 & -4\end{pmatrix}$. Find the $(1,2)$ and $(3,2)$ cofactors of A.

- $C_{1,2}(A)=(-1)^{1+2}\times\det(A_{1,2})=-\begin{vmatrix}4 & -1\\ 0 & -4\end{vmatrix}=16$.
- $C_{3,2}(A)=(-1)^{3+2}\times\det(A_{3,2})=-\begin{vmatrix}1 & 3\\ 4 & -1\end{vmatrix}=13$.

Definition

- For $n=1,2$, the determinant of $n\times n$ matrices is defined.
- Assume we have defined the determinant for $(n-1)\times(n-1)$ matrices. We want to define it for $n\times n$ matrices.
- Let A be an $n\times n$ matrices. Let $C_{ij}(A)$ be its cofactors. We define $\det A=a_{1,1}C_{1,1}(A)+a_{2,1}C_{2,1}(A)+\cdots+a_{n,1}C_{n,1}(A)$. This is the <u>Laplace Expansion</u>.

Example

Let $A=\begin{pmatrix}1 & 3 & -1\\ 0 & -1 & 2\\ 2 & 4 & 0\end{pmatrix}$. Compute $\det A$.

- $\det A=1\times C_{1,1}(A)+0\times C_{2,1}(A)+2\times C_{3,1}(A)$.
- $C_{1,1}(A)=(-1)^{1+1}\begin{vmatrix}-1 & 2\\ 4 & 0\end{vmatrix}=-8$. $C_{3,1}(A)=(-1)^{3+1}\begin{vmatrix}3 & -1\\ -1 & 2\end{vmatrix}=5$.
- So, $\det A=-8+2(5)=2$.

PROPERTIES OF THE DETERMINANT

Theorem

Let $A=a_{ij}$ be an $n\times n$ matrix. Then the determinant of A is given by:

- $\det A=a_{i,1}C_{i,1}(A)+a_{i,2}C_{i,2}(A)+\cdots+a_{i,n}C_{i,n}(A), i=1,\cdots,n$ (expansion along row i).

- $\det A = a_{1,j}C_{1,j}(A) + a_{2,j}C_{2,j}(A) + \cdots + a_{n,j}C_{n,j}(A),\ j = 1, \cdots, n$ (expansion along row *j*).

Example

Let $A = \begin{pmatrix} 1 & 0 & 1- & 4 \\ 3 & 0 & 3 & 5 \\ 1 & -1 & 2 & -2 \\ 4 & 0 & -3 & 0 \end{pmatrix}$. What is $\det A$?

$\det A = a_{1,2}C_{1,2}(A) + a_{2,2}C_{2,2}(A) + a_{3,2}C_{3,2}(A) + a_{4,2}C_{4,2}(A) = -C_{3,2}(A)$

- $= -(-1)^{3+2}\begin{vmatrix} 1 & 1 & -4 \\ 3 & 3 & 5 \\ 4 & -3 & 0 \end{vmatrix} = \begin{vmatrix} 1 & 1 & -4 \\ 3 & 3 & 5 \\ 4 & -3 & 0 \end{vmatrix} =: A_1$.
- $\det(A_1) = -4C_{1,3}(A_1) + 5C_{2,3}(A_1) = -4(-1)^{1+3}\begin{vmatrix} 3 & 3 \\ 4 & -3 \end{vmatrix} + 5(-1)^{2+3}\begin{vmatrix} 1 & 1 \\ 4 & -3 \end{vmatrix} = -4(-21) - 5(-7) = 119$.

Theorem

Let *A* be an *n*×*n* matrix. Then:

1) If *A* has a row (column) of 0's, then $\det A = 0$.
2) If we interchange two rows (columns), then the resulting determinant is $-\det A$. So $\det\begin{pmatrix} C_1 & \cdots & C_i & C_j & \cdots & C_n \end{pmatrix} = -\det\begin{pmatrix} C_1 & \cdots & C_j & C_i & \cdots & C_n \end{pmatrix}$.
3) If a row (column) of *A* is multiplied by $u \in \mathbf{R}$, then the determinant is multiplied by *u*. So $\det\begin{pmatrix} C_1 & \cdots & uC_i & \cdots & C_n \end{pmatrix} = u\det\begin{pmatrix} C_1 & \cdots & C_i & \cdots & C_n \end{pmatrix}$.
4) If two rows (columns) of *A* are equal, then $\det A = 0$.
5) If a multiple of a row *i* (column *j*) is added to a different row (column), then the determinant is the same. So $\det\begin{pmatrix} C_1 & \cdots & C_i & C_j & \cdots & C_n \end{pmatrix} = \det\begin{pmatrix} C_1 & \cdots & C_i & C_j + aC_i & \cdots & C_n \end{pmatrix}$.

Examples

1) Compute $\det\begin{pmatrix} 1 & 1 & 2 \\ -3 & 0 & 3 \\ 2 & -1 & -4 \end{pmatrix} = -3(-1)^{2+1}\begin{vmatrix} 1 & 3 \\ -1 & -2 \end{vmatrix} = 3(1) = 3$.
2) Compute $\det\begin{pmatrix} 1 & 100 & 0 \\ 2 & -1 & 0 \\ 15 & 3 & 0 \end{pmatrix} = 0$.

Example

Let $A = \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ {x_1}^2 & {x_2}^2 & {x_3}^2 \end{pmatrix}, x_1, x_2, x_3, x_4 \in \mathbf{R}$. What is $\det A$?

$$\det A = \begin{vmatrix} 1 & 0 & 0 \\ x_1 & x_2 - x_1 & x_3 - x_1 \\ {x_1}^2 & {x_2}^2 - {x_1}^2 & {x_3}^2 - {x_1}^2 \end{vmatrix} = C_{1,1}(A) = (-1)^{1+1} \begin{vmatrix} x_2 - x_1 & x_3 - x_1 \\ {x_2}^2 - {x_1}^2 & {x_3}^2 - {x_1}^2 \end{vmatrix}.$$

$$= (x_2 - x_1)(x_3 - x_1) \begin{vmatrix} 1 & 1 \\ x_2 + x_1 & x_3 + x_1 \end{vmatrix} = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2)$$

Corollary

A is an $n \times n$ matrix. Then $\det(uA) = u^n \det A, u \in \mathbf{R}$.

Theorem

Let A be a triangle matrix (i.e.: $a_{ij} = 0$ when $i > j$ or $i < j$). Then $\det A$ is the product of the entries of the main diagonal of A. So $\det A = a_{1,1} \times a_{2,2} \times \cdots \times a_{n,n}$

Example

Let $A = \begin{pmatrix} 4 & 3 & 2 \\ 3 & -2 & 5 \\ 2 & 4 & 6 \end{pmatrix}$. What is $\det A$?

$$\det A = 2\begin{vmatrix} 4 & 3 & 2 \\ 3 & -2 & 5 \\ 1 & 2 & 3 \end{vmatrix} = -2\begin{vmatrix} 1 & 2 & 3 \\ 3 & -2 & 5 \\ 4 & 3 & 2 \end{vmatrix} = -2\begin{vmatrix} 1 & 2 & 3 \\ 0 & -8 & -4 \\ 0 & -5 & -10 \end{vmatrix} = (-2)(-4)(-5)\begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{vmatrix} = 40\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{vmatrix}$$

$$= 40\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -3 \end{vmatrix} = (40)(-3) = -120$$

Theorem

Consider $C = \begin{bmatrix} A & X \\ 0 & B \end{bmatrix}$, $D = \begin{bmatrix} A & 0 \\ Y & B \end{bmatrix}$ for A, B square matrices. Then:

- $\det C = \det A \times \det B$.
- $\det D = \det A \times \det B$.

Example

Let $C = \begin{pmatrix} 1 & 3 & 2 & -1 \\ 2 & -1 & 3 & 4 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 1 & 5 \end{pmatrix}$. What is $\det C$?

- $C = \left(\begin{array}{cc|cc} 1 & 3 & 2 & -1 \\ 2 & -1 & 3 & 4 \\ \hline 0 & 0 & -2 & 1 \\ 0 & 0 & 1 & 5 \end{array}\right)$. So $\det C = \begin{vmatrix} 1 & 3 \\ 2 & -1 \end{vmatrix} \times \begin{vmatrix} -2 & 1 \\ 1 & 5 \end{vmatrix} = (-7)(-11) = 77$.

DETERMINANTS OF PRODUCT OF MATRICES

Theorem

Let A, B be $n \times n$ matrices. Then $\det(AB) = \det A \times \det B$.

- Remark: If A, B are $n \times n$ matrices, $\det(A+B) \neq \det A + \det B$.

Corollary

Let A be an $n \times n$ matrix. Then $\det(A^k) = (\det A)^k$.

Proof (by induction):

- Prove for $k = 1$:
  - $\det(A^1) = \det A = (\det A)^1$. The corollary is true for $k = 1$.
- Assume $\det(A^k) = (\det A)^k$. Prove $\det(A^{k+1}) = (\det A)^{k+1}$.
  - $\det(A^{k+1}) = \det(A^k A) = \det(A^k) \times \det A = (\det A)^k (\det A) = (\det A)^{k+1}$. So the corollary is true for $k+1$.

Lecture #18 – Thursday, March 11, 2004

DETERMINANTS AND MATRIX INVERSES

Theorem

Let A be an $n \times n$ matrix. A is invertible iff $\det A \neq 0$.

Corollary

If A is an $n \times n$ matrix that is invertible. Then $\det(A^{-1}) = \dfrac{1}{\det A}$.

Proof:

- Remark $\det(I_n) = 1$, and $A^{-1}A = I_n$.
- $\det(A^{-1}A) = \det(I_n) = 1 \Rightarrow (\det A^{-1})(\det A) = 1 \Rightarrow \det A^{-1} = \dfrac{1}{\det A}$.

Example

For which values of $a \in \mathbf{R}$ is the matrix $A = \begin{pmatrix} 1 & 0 & -a \\ 2 & 3 & 1 \\ a & 2a & 1 \end{pmatrix}$ invertible?

- $\det A = \begin{vmatrix} 1 & 0 & 0 \\ 2 & 3 & 1+2a \\ a & 2a & 1+a^2 \end{vmatrix} = 1 \times C_{11}(\tilde{A}) = (-1)^{1+1} \begin{vmatrix} 3 & 1+2a \\ 2a & 1+a^2 \end{vmatrix} = 3(1+a^2) - (2a)(1+2a) = -(a^2+2a-3)$.
- A is invertible when $-(a^2+2a-3) \neq 0$. $\Delta = (2)^2 - (4)(-3) = 16 = 4^2$, so $a_1 = -3$, $a_2 = 1$.

- So A is invertible if a is different from -3 and 1.

Proposition

Let A be an $n\times n$ matrix. Then $\det(A)=\det(A^T)$.

Example

Let A, B be square matrices. If $\det A=2$, $\det B=-3$, compute $\det(A^2B^TA^{-1}B^3)$.

- $\det(A^2B^TA^{-1}B^3)=(\det A)^2(\det B)\left(\frac{1}{\det A}\right)(\det B)^4=(2)(-3)^5=162$.

Definition

A square matrix A is orthogonal iff $A^{-1}=A^T$ (so A is invertible).

Proposition

If A is orthogonal, then $\det A=\pm 1$.

Proof: $AA^{-1}=I \Leftrightarrow AA^T=I \Rightarrow \det(AA^T)=\det(I)=1 \Leftrightarrow (\det A)(\det A^T)=1 \Leftrightarrow (\det A)^2=1 \Leftrightarrow \det A=\pm 1$.

Matrix Adjoint

Definition

Let A be an $n\times n$ matrix. Let $C_{ij}(A)$ be the cofactors of A. The adjoint of A is $\operatorname{adj} A=[C_{ij}(A)]^T$.

Theorem

Let A be an $n\times n$ matrix. Then $A(\operatorname{adj} A)=(\det A)I=(\operatorname{adj} A)A$. In particular, if $\det A\neq 0$, then $A^{-1}=\frac{1}{\det A}\operatorname{adj}(A)$.

Proof: $A\times\operatorname{adj} A=(\det A)I \Rightarrow A\left(\frac{1}{\det A}\operatorname{adj} A\right)=I$, so $A^{-1}=\frac{1}{\det A}\operatorname{adj}(A)$.

Corollary

Let A be an $n\times n$ matrix. If $\det A\neq 0$, then $\det(\operatorname{adj} A)=(\det A)^{n-1}$.

Proof:
$(A)(\operatorname{adj} A)=(\det A)I \Rightarrow \det(A\operatorname{adj} A)=\det((\det A)I) \Rightarrow (\det A)(\det(\operatorname{adj} A))=(\det A)^n \Rightarrow \det(\operatorname{adj} A)=(\det A)^{n-1}$.

Example

Let $A=\begin{pmatrix}1 & 1 & a\\ -a & 1 & -a\\ a & -1 & 1\end{pmatrix}$. For what values of a is A invertible? When it is the case, compute A^{-1}.

- $\det A = \begin{vmatrix} 1 & 0 & 0 \\ -a & 1+a & -a+a^2 \\ a & -1-a & 1-a^2 \end{vmatrix} = 1 \times (-1)^{1+1} \begin{vmatrix} 1+a & -a+a^2 \\ -1-a & 1-a^2 \end{vmatrix} = 1+a^2$
- So A is invertible when $a \neq \pm 1$.
- $(\operatorname{adj} A)^T = \begin{pmatrix} C_{11}(A) & C_{12}(A) & C_{13}(A) \\ C_{21}(A) & C_{22}(A) & C_{23}(A) \\ C_{31}(A) & C_{32}(A) & C_{33}(A) \end{pmatrix}$, and $[C_{ij}(A)]^T = \operatorname{adj} A = \begin{pmatrix} 1-a & -1-a & -2a \\ a-a^2 & 1-a^2 & a-a^2 \\ 0 & 1+a & 1+a \end{pmatrix}$.
- When $a \neq \pm 1$, $A^{-1} = \dfrac{1}{\det A} \operatorname{adj} A = \dfrac{1}{1-a^2} \begin{pmatrix} 1-a & -1-a & -2a \\ a-a^2 & 1-a^2 & a-a^2 \\ 0 & 1+a & 1+a \end{pmatrix}$.

Cramer's Rule

Theorem: Cramer's Rule

If A is an invertible matrix, then the solution of the system $AX = B$ is $X = A^{-1}B$. If $A = [C_1 \ \cdots \ C_n], C_j \in \mathbf{R}^n$ and $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, consider $A_j = (C_1 \ \cdots \ C_{j-1} \ B \ C_{j+1} \ C_n)$ Then $x_j = \dfrac{\det(A_j)}{\det(A)}$.

Example

Solve x_1 for $\begin{cases} 3x_1 + 2x_2 + x_3 = 2 \\ x_1 + x_2 + 2x_3 = 0 \\ -x_1 + 2x_2 + x_3 = 1 \end{cases}$.

- Let $A = \begin{pmatrix} 3 & 2 & 1 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$, $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Then $AX = B$.
- $\det(A) = -12$. So $x_1 = \dfrac{\det(A_1)}{\det(A)} = \dfrac{\begin{vmatrix} 2 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 1 \end{vmatrix}}{\det A} = \dfrac{-3}{-12} = \dfrac{1}{4}$.

Lecture #19 – Tuesday, March 16, 2004

Permutations

Definition

Let $S_n = \{1, 2, \cdots, n\}$. A map $\sigma : S_n \to S_n$ is a <u>permutation</u> if it is a one-to-one map (rearrangement).

Examples

1) If $S_3 = \{1,2,3\}$, a possible permutation is $\begin{matrix}\sigma_1(1)=1\\ \sigma_1(2)=2\\ \sigma_1(3)=3\end{matrix}$. Another possible permutation is $\begin{matrix}\sigma_2(1)=2\\ \sigma_2(2)=3\\ \sigma_2(3)=1\end{matrix}$.

2) If $S_2 = \{1,2\}$, the possible permutations are $\begin{matrix}\sigma_1(1)=1\\ \sigma_1(2)=2\end{matrix}$ and $\begin{matrix}\sigma_2(1)=2\\ \sigma_2(2)=1\end{matrix}$.

Definition

Let $\sigma : S_n \to S_n$ be a permutation. σ has an inversion if there exists $i < j$ in S_n such that $\sigma(i) > \sigma(j)$.

- We say σ is even if the number of inversions is even.
- We say σ is odd if the number of inversions is odd.

Example

If $S_2 = \{1,2\}$, then all possible permutations are $\begin{matrix}\sigma_1(1)=1\\ \sigma_1(2)=2\end{matrix}$ and $\begin{matrix}\sigma_2(1)=2\\ \sigma_2(2)=1\end{matrix}$.

- Since σ_1 has 0 inversions, σ_1 is even.
- Since σ_2 has 1 inversion, σ_2 is odd.

Example

Let $S_3 = \{1,2,3\}$ and $\begin{matrix}\sigma(1)=2\\ \sigma(2)=1\\ \sigma(3)=3\end{matrix}$. Since $1 < 2$ and $\sigma(1) > \sigma(2)$, σ has 1 inversion, and so σ is odd.

PERMUTATION AND DETERMINANTS

Proposition

Let A be an $n \times n$ matrix. If $A = (a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$, then $\det A = \sum_{\sigma : s_n \to s_n} (\pm 1)\left(a_{1,\sigma(1)} \times a_{2,\sigma(2)} \times \cdots \times a_{n,\sigma(n)}\right)$

and + if σ is even, and - if σ is odd.

Example

Let $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $S_2 = \{1,2\}$. Find $\det A$.

- The possible permutations are $\begin{matrix}\sigma_1(1)=1\\ \sigma_1(2)=2\end{matrix}$ (even) and $\begin{matrix}\sigma_2(1)=2\\ \sigma_2(2)=1\end{matrix}$ (odd).
- So, $\det A = \sum_{\sigma = \{\sigma_1, \sigma_2\}} (\pm 1)\left(a_{1,\sigma(1)} \times a_{2,\sigma(2)}\right) = a_{1,\sigma_1(1)} \times a_{2,\sigma_1(2)} - a_{1,\sigma_2(1)} \times a_{2,\sigma_2(2)} = a_{11}a_{22} - a_{12}a_{21}$.

Proposition

Let $A = (a_{ij})$ be an $n \times n$ matrix.

1) If A has a row (column) of 0, then $\det A = 0$.
2) If $\tilde{A}$ is formed from A where R_j was replaced by uR_j for $u \in \mathbf{R}$, then $\det A = u \det \tilde{A}$.
3) If $\tilde{A}$ is formed from A where we exchange R_k and R_l, then $\det A = \det \tilde{A}$.

Proof of (1): We have $\det A = \sum_{\sigma : s_n \to s_n} (\pm 1)\left(a_{1,\sigma(1)} \times \cdots \times a_{n,\sigma(n)}\right)$.

- Suppose R_j has only zero entries, that is $a_{jk} = 0, k = 1, 2, \cdots, n$. Then
  $\det A = \sum_{\sigma : s_n \to s_n} (\pm 1)\left(a_{1,\sigma(1)} \times \cdots \times a_{j,\sigma(j)} \times \cdots \times a_{n,\sigma(n)}\right) = 0$ since $a_{j,\sigma(j)} = 0$.

Proof of (2):

- If $A = \begin{pmatrix} R_1 \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix}$, $\tilde{A} = \begin{pmatrix} R_1 \\ \vdots \\ uR_j \\ \vdots \\ R_n \end{pmatrix}$, where $\tilde{A} = (\tilde{a}_{kl}) = (a_{kl})$ for $k \neq l$ and $\tilde{a}_{kl} = ua_{kl}, l = 1, \cdots, n$ for $k = l$.
- $$\begin{aligned}\det \tilde{A} &= \sum_{\sigma : s_n \to s_n} (\pm 1)\left(\tilde{a}_{1,\sigma(1)} \times \cdots \times \tilde{a}_{j,\sigma(j)} \times \cdots \times \tilde{a}_{n,\sigma(n)}\right) = \sum_{\sigma : s_n \to s_n} (\pm 1)\left(a_{1,\sigma(1)} \times \cdots \times ua_{j,\sigma(j)} \times \cdots \times a_{n,\sigma(n)}\right) \\ &= u \sum_{\sigma : s_n \to s_n} (\pm 1)\left(a_{1,\sigma(1)} \times \cdots \times a_{j,\sigma(j)} \times \cdots \times a_{n,\sigma(n)}\right) = u \det A\end{aligned}.$$

Example

Compute $D = \begin{vmatrix} 1 & \cos a & \cos 2a \\ \cos a & \cos 2a & \cos 3a \\ \cos 2a & \cos 3a & \cos 4a \end{vmatrix}, a \in \mathbf{R}$.

- $$\begin{aligned} D &= \begin{vmatrix} 1 & 0 & 0 \\ \cos a & \cos 2a - \cos^2 a & \cos 3a - \cos 2a \cos a \\ \cos 2a & \cos 3a - \cos 2a \cos a & \cos 4a - \cos^2 2a \end{vmatrix} = (-1)^{1+1} \begin{vmatrix} -\sin^2 a & -\sin 2a \sin a \\ -\sin 2a \sin a & -\sin^2 2a \end{vmatrix}. \\ &= (\sin^2 2a)(\sin^2 a) - (\sin 2a \sin a)^2 = 0 \end{aligned}$$

POWER OF MATRICES

- Remark: A^2 exists only if A is square.

Definition

A matrix A is diagonal if all the entries are zero except the ones on the main diagonal.

Examples

1) $A = I_n$ is a diagonal matrix.
2) $A = 0$ is a diagonal matrix.

3) $A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$ is a diagonal matrix.

Example

Let $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \lambda_1, \lambda_2 \in \mathbf{R}$. Prove that $A^k = \begin{pmatrix} \lambda_1{}^k & 0 \\ 0 & \lambda_2{}^k \end{pmatrix}, k \geq 1 \in \mathbf{R}$.

- Prove for $k = 1$:
  - $A^1 = A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1{}^1 & 0 \\ 0 & \lambda_2{}^1 \end{pmatrix}$. So it is true for $k = 1$.
- Assume $A^k = \begin{pmatrix} \lambda_1{}^k & 0 \\ 0 & \lambda_2{}^k \end{pmatrix}$. Prove $A^{k+1} = \begin{pmatrix} \lambda_1{}^{k+1} & 0 \\ 0 & \lambda_2{}^{k+1} \end{pmatrix}$:
  - $A^{k+1} = A^k A = \begin{pmatrix} \lambda_1{}^k & 0 \\ 0 & \lambda_2{}^k \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1{}^{k+1} & 0 \\ 0 & \lambda_2{}^{k+1} \end{pmatrix}$. So it is true for $k+1$.

Proposition

Let A be a diagonal matrix. $A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} := \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then

$A^k = \operatorname{diag}(\lambda_1{}^k, \lambda_2{}^k, \ldots, \lambda_n{}^k) = \begin{pmatrix} \lambda_1{}^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n{}^k \end{pmatrix}$ for $k \geq 1$.

Corollary

Let P be a polynomial. $P(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$. We define $P(A) = a_m A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 I$. If $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, then $P(A) = \operatorname{diag}(P(\lambda_1), \ldots, P(\lambda_n))$.

DIAGONALIZATION

Definition

Let A be an $n \times n$ matrix. We say A is <u>diagonalizable</u> if there exists an <u>invertible matrix</u> P such that $P^{-1}AP$ is a diagonal matrix. P is called the <u>diagonalizing matrix</u>.

Theorem

Let A be a diagonalizable matrix. Let D be a diagonal matrix. If there exists P an invertible matrix such that $P^{-1}AP = D \Rightarrow A = PDP^{-1}$, then $A^k = PD^k P^{-1}$.

Proof by Induction:

- For $k = 1$, A is diagonalizable, so $A = PDP^{-1}$.
- Assume $A^k = PD^k P^{-1}$. Prove $A^{k+1} = PD^{k+1}P^{-1}$:
  - $A^{k+1} = A^k A = PD^k P^{-1}PDP^{-1} = PD^k \left(P^{-1}P\right)DP^{-1} = PD^k DP^{-1} = PD^{k+!}P^{-1}$.

Example

Let $A = \begin{pmatrix} -1 & 0 \\ -3 & -2 \end{pmatrix}$, $P = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Computation gives $P^{-1} = \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}$. Caculate A^k.

- $P^{-1}AP = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} = D$. So $A^k = PD^kP^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2^k & 0 \\ 0 & (-1)^k \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} (-1)^k & 0 \\ (-1)^k - 2^k & 2^k \end{pmatrix}$.

EIGENVALUES AND EIGENVECTORS

Definition

Let A be an $n\times n$ matrix. $\lambda \in \mathbf{R}$ is an eigenvalue of A if there exists some vector $X \neq 0$ such that $AX = \lambda X$. X is an eigenvector of A for λ. X is also called a λ-eigenvector.A9

- Remark: If $X = 0$, $AX = 0 = \lambda 0$ for all $\lambda \in \mathbf{R}$. This is meaningless!
- Remark: Let X be a λ-eigenvector ($X \neq 0$) for A, and $a \neq 0 \in \mathbf{R}$. Then $a(X) \neq 0$ is also a λ-eigenvector.
  Proof: $A(aX) = a(AX) = a\lambda X = \lambda(aX)$.

Proposition

Let A be an $n\times n$ matrix. Let λ be an eigenvalue of A. Let X be a λ-eigenvector. Then λ^k is an eigenvalue of A^k, and X is a λ^k-eigenvector for $k \geq 1$.

Proof by Induction:

- For $k = 1$, it is clear.
- Assume λ^k is an eigenvalue of A^k and X is a λ^k-eigenvector. Prove for $k+1$:
  - $A^{k+1}X = A^k(AX) = A^k(\lambda X) = \lambda A^k X = \lambda\lambda(^k X) = \lambda^{k+1}X$.

Example

Consider $A = \begin{pmatrix} -1 & 0 \\ -3 & -2 \end{pmatrix}$.

- $A\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, so 2 is an eigenvalue of A and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is a 2-eigenvector of A.
- $A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix} = -1\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, so -1 is an eigenvalue of A and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is a -1-eigenvector of A.
- Notice: $P = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ such that $P^{-1}AP = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$.

Lecture #20 – Thursday, March 18, 2004

Proposition

0 is an eigenvalue of A iff A is not invertible.

Proof:

- Assume 0 is an eigenvalue of A. $AX = 0X = 0$. Since there exists non-trivial solutions, so A is not invertible.
- Assume A is not invertible. There exists $X \neq 0$ such that $AX = 0X$. So 0 is an eigenvalue of A.

Definition

Let A be an $n\times n$ matrix. Let $c_A(x) = \det(xI - A), x \in \mathbf{R}$. c_A is a polynomial of degree n. c_A is the characteristic polynomial of A.

Theorem

1) The eigenvalue of A are the roots of $c_A(x)$. That is λ is an eigenvalue of A iff $c_A(\lambda) = 0$.
2) The λ-eigenvectors X are the non-trivial solutions of the homogeneous system $(\lambda I - A)X = 0$.

- Proof (1): Let λ be an eigenvalue of A. So $\exists$ $X \neq 0$ such that $AX = \lambda X \Leftrightarrow (\lambda I - A)X = 0$. Since $\lambda I - A$ is not invertible, $\det(\lambda I - A) = 0 = c_A(\lambda)$. So $c_A(\lambda) = 0$.
- Proof (2): Let λ such that $c_A(\lambda) = 0 \Leftrightarrow \det(\lambda I - A) = 0$. So $\lambda I - A$ is not invertible. So $\exists$ $X \neq 0$ such that $(\lambda I - A)X = 0 \Leftrightarrow \lambda X - AX = 0 \Leftrightarrow \lambda X = AX$. So λ is an eigenvalue.

Example

Let $A = \begin{pmatrix} -5 & 6 \\ -9 & 10 \end{pmatrix}$. What is $c_A(x)$? Find the λ-eigenvectors.

- $c_A(x) = \det(xI - A) = \left| \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} - \begin{pmatrix} -5 & 6 \\ -9 & 10 \end{pmatrix} \right| = \begin{vmatrix} x+5 & -6 \\ 9 & x-10 \end{vmatrix} = x^2 - 5x + 4 = (x-4)(x-1)$.
- So the eigenvectors of A are 4 and 1.
- Want: The 4-eigenvectors:
  - X is a 4-eigenvector if

    $$AX = 4X \Leftrightarrow (4I - A)X = 0 \Leftrightarrow \begin{pmatrix} 9 & -6 \\ 9 & -6 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \Leftrightarrow \begin{cases} 9x_1 - 6x_2 = 0 \\ 9x_1 - 6x_2 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = \frac{2}{3}t \\ x_2 = t \end{cases}, t \in \mathbf{R}.$$

  - So $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = t\begin{pmatrix} \frac{2}{3} \\ 1 \end{pmatrix}$ is an eigenvector for $t \neq 0$.
- Want: The 1-eigenvectors:
  - X is a 1-eigenvector if

    $$(I - A)X = 0 \Leftrightarrow \begin{pmatrix} 6 & -6 \\ 9 & -9 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \Leftrightarrow \begin{cases} 6x_1 - 6x_2 = 0 \\ 9x_1 - 9x_2 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = t \\ x_2 = t \end{cases}, t \in \mathbf{R}.$$

  - So $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = t\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is an eigenvector for $t \neq 0$.

Proposition

Let A be an $n\times n$ matrix. Let λ be an eigenvalue of A. Let $X_1, X_1, \ldots, X_k$ be some λ-eigenvectors. Then any linear combinations of X_j that is not 0 is a λ-eigenvector.

Proof: Let $X = a_1X_1 + a_2X_2 + \cdots + a_kX_k \neq 0$. Since $AX_j = \lambda X_j$, $j = 1,2,\ldots,k$,
$AX = A(a_1X_1 + \cdots + a_kX_k) = a_1AX_1 + \cdots + a_kAX_k = a_1\lambda X_1 + \cdots + a_k\lambda X_k = \lambda(a_1X_1 + \cdots + a_kX_k) = \lambda X$. So $X = a_1X_1 + a_2X_2 + \cdots + a_kX_k \neq 0$ is a λ-eigenvector.

Example

Let $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Show $A^2 = 2I + A$. What are the possible eigenvalues of *A*?

- $A^2 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} = 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = 2I + A$.
- Let λ be an eigenvalue of *A*. Let *X* be a λ-eigenvector. So $AX = \lambda X$.
- $A^2X = AAX = A(AX) = A\lambda X = \lambda AX = \lambda^2 X$ and
  $A^2X = (2I + A)X = 2X + AX = 2X + \lambda X = (2 + \lambda)X$.
- So $\lambda^2 X = (2+\lambda)X \Rightarrow \lambda^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (2+\lambda)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Since $X \neq 0$, so suppose $x_1 \neq 0$. Then
  $\lambda^2 x_1 = (2+\lambda)x_1 \Rightarrow \lambda^2 = 2 + \lambda$
- If λ is an eigenvalue of *A*, then $\lambda^2 - \lambda - 2 = 0 \Leftrightarrow (\lambda - 2)(\lambda + 1) = 0$. So the possible eigenvalues of *A* are -1 and 2.
- For $\lambda = -1$, the -1-eigenvector $X = s\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, (s,t) \neq (0,0)$.
- For $\lambda = 2$, the 2-eigenvector $X = t\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, t \neq 0$.

Lecture #21 – Tuesday, March 23, 2004

EIGENSPACES

Note: *A* is an $n \times n$ matrix.

Definition

Let λ be an eigenvalue of *A*. Then the λ-eigenspace is defined by $E_\lambda(A) = \text{null}(\lambda I - A) = \{X \in \mathbf{R}, AX = \lambda X\}$.

Proposition

Let $X_1, \ldots, X_k$ be the basic solution of $(\lambda I - A)X = 0$. Then $E_\lambda(A) = \text{span}(X_1, \ldots, X_k)$. Moreover, $\{X_1, \ldots, X_k\}$ is a basis of the subspace $E_\lambda(A)$. ($\dim(E_\lambda(A)) = k$)

Example

What are the eigenspaces of $A=\begin{pmatrix}2&3\\0&2\end{pmatrix}$?

- Eigenvalues: $c_A(x)=\det(xI-A)=\begin{vmatrix}x-2&-3\\0&x-2\end{vmatrix}=(x-2)^2$. So $\lambda=2$.
- $E_2(A)$: $(2I-A)X=\begin{pmatrix}0&-3\\0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}\Rightarrow\begin{cases}-3x_2=0\\0=0\end{cases}\Rightarrow X=\begin{pmatrix}s\\0\end{pmatrix}=s\begin{pmatrix}1\\0\end{pmatrix}$. So $E_2(A)=\left\{X=s\begin{pmatrix}1\\0\end{pmatrix},s\in\mathbf{R}\right\}$.

Proposition

Let $\lambda\neq\mu$ be two eigenvalues of A. Then $E_\lambda(A)\cap E_\mu(A)=\{0\}$. ($E_\lambda(A)\oplus E_\mu(A)$)

Proof: Let $X\in E_\lambda(A)\cap E_\mu(A)$. So $\left.\begin{matrix}AX=\lambda X\\AX=\mu X\end{matrix}\right\}\Rightarrow(\lambda-\mu)X=0\Rightarrow X=0$.

Theorem

Let $\lambda_1,\ldots,\lambda_k$ be distinct eigenvalues of A. Let $X_1,\ldots,X_k$ be the eigenvectors of A. Then $\{X_1,\ldots,X_k\}$ are linearly independent.

Proof (by induction):

- $k=2$:
  - $\alpha_1X_1+\alpha_2X_2=0,\alpha_1,\alpha_2\in\mathbf{R}\Rightarrow\alpha_1X_1=-\alpha_2X_2$.
  - Since $\alpha_1X_1\in E_{\lambda_1}(A)$ and $\alpha_2X_2\in E_{\lambda_2}(A)$, and $E_{\lambda_2}(A)\cap E_{\lambda_2}(A)=\{0\}$. So $\alpha_1X_1=0$, and $\alpha_2X_2=0$, so $\alpha_1=\alpha_2=0$.
- Assume $X_1,\ldots,X_k$ are linearly independent. We want to prove $X_1,\ldots,X_k,X_{k+1}$ are linearly independent.
  - $\alpha_1X_1+\cdots+\alpha_kX_k+\alpha_{k+1}X_{k+1}=0$ L_1.
  - $A(\alpha_1X_1+\cdots+\alpha_kX_k+\alpha_{k+1}X_{k+1})=0\Rightarrow\alpha_1\lambda_1X_1+\cdots+\alpha_k\lambda_kX_k+\alpha_{k+1}\lambda_{k+1}X_{k+1}=0$ L_2.
  - So, $L_1-\lambda_{k+1}L_2=\alpha_1(\lambda_1-\lambda_{k+1})X_1+\cdots\alpha_k(\lambda_1-\lambda_{k+1})X_k=0$.
  - Since $X_1,\ldots,X_k$ are linearly independent, $\left.\begin{matrix}\alpha_1(\lambda_1-\lambda_{k+1})=0\\\vdots\\\alpha_k(\lambda_1-\lambda_{k+1})=0\end{matrix}\right\}\Rightarrow\alpha_1=\cdots=\alpha_k=0$.
  - So, from L_1, we have $\alpha_{k+1}X_{k+1}=0\Rightarrow\alpha_{k+1}=0$.

Example

Let $A=\begin{pmatrix}1&-3&3\\-1&-1&1\\2&-2&2\end{pmatrix}$. Show that there is a basis of $\mathbf{R}^3$ of eigenvectors.

- Eigenvalues: $c_A(x)=x(x+2)(x-4)$, so $\lambda_1=0$, $\lambda_2=-2$, $\lambda_3=4$.
- So since X_1 is the 0-eigenvector, X_2 the -2-eigenvector, X_3 the 4-eigenvector, and they corresponds to 3 distinct, linearly independent eigenvectors, we have a basis. ($\dim\mathbf{R}^3=3$)

Example

Let $A = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}$. Prove that there is no basis of eigenvectors.

- Eigenvalues: $c_A(x) = (x-2)^2$. So $\lambda = 2$.
- 2-eigenvectors: $E_2(A) = \left\{ X \in \mathbf{R}^2, X = s\begin{pmatrix} 1 \\ 0 \end{pmatrix}, s \in \mathbf{R} \right\}$.
- Since $\dim(E_2(A)) = 1$, there is no basis of eigenvectors.

DIAGONALIZATION OF MATRICES: SIMILAR MATRICES

Definition

Let *A*, *B* be two *n*×*n* matrices. We say that *A* and *B* are similar if there exists an invertible matrix *P* such that $B = P^{-1}AP$. ($B \sim A$)

Remark

We can rewrite the definition of diagonalizable matrices: *A* is diagonalizable if there exists *D* a diagonal matrix such that *D* and *A* are similar. ($A \sim D$)

Proposition

1) We have $A \sim A$.
2) If $A \sim B$, then $B \sim A$.
3) If $A \sim B$ and $B \sim C$, then $A \sim C$.

Proof:

1) $A = I^{-1}AI = IAI = A$.
2) $B = P^{-1}AP$, *P* is invertible. Let $Q = P^{-1}$, then $A = Q^{-1}BQ = PBP^{-1}$.
3) $A \sim B$ means $A = P^{-1}BP$. $B \sim C$ means $B = P^{-1}CP$. So
$A = P^{-1}BP = P^{-1}\left(Q^{-1}CQ\right)P \Leftrightarrow \left(P^{-1}Q^{-1}\right)C(QP) \Leftrightarrow (QP)^{-1}C(QP)$, so $A \sim C$.

Corollary

If *A* is diagonalizable, and $B \sim A$, then *B* is also diagonalizable.

Proof: *A* is diagonalizable. Then there exists *D* diagonal matrix such that $A \sim D$. $B \sim A \sim D \Rightarrow B \sim D$.

Proposition

If $A \sim B$, then:

1) $A^{-1} \sim B^{-1}$.
2) $A^T \sim B^T$.
3) For any $k \in \mathbf{N}$, $A^k \sim B^k$.

Proof:

1) $A \sim B \Rightarrow A = P^{-1}BP$. So $A^{-1} = \left(P^{-1}BP\right)^{-1} = \left(P^{-1}\right)\left(B^{-1}\right)\left(P^{-1}\right)^{-1} = P^{-1}B^{-1}P \Rightarrow A^{-1} \sim B^{-1}$.
2) $A^T = \left(P^{-1}BP\right)^T = P^T B^T \left(P^{-1}\right)^T = P^T B^T \left(P^T\right)^{-1}$.

Definition

Let A be an $n\times n$ matrix, where $A=(a_{ij})$. The trace of A is defined by $\operatorname{tr}(A)=\sum_{j=}^{n} ajj = a_{11}+a_{22}+\cdots+a_{nn}$.

Proposition

We have:

1) $\operatorname{tr}(A+B)=\operatorname{tr}(A)+\operatorname{tr}(B)$.
2) If $u\in\mathbf{R}$, $\operatorname{tr}(uA)=u\operatorname{tr}(A)$.
3) $\operatorname{tr}(AB)=\operatorname{tr}(BA)$.

Proof:

1) It is trivial.
2) It is trivial.
3) Let $A=(a_{ij})$, $B=(b_{ij})$, $C=AB=(c_{ij})$, $D=BA=(d_{ij})$.

- $a_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}=a_{i1}b_{1i}+\cdots+a_{in}b_{nj}$.
- $d_{ij}=\sum_{k=1}^{n} b_{jk}a_{kj}$.
- $\operatorname{tr}(AB)=\sum_{l=1}^{n} c_{ll}=\sum_{l=1}^{n}\left(\sum_{k=1}^{n} a_{lk}b_{kl}\right)=\sum_{k,l=1}^{n} a_{lk}b_{kl}=\sum_{k,l=1}^{n} b_{kl}a_{lk}=\sum_{k=1}^{n}\left(\sum_{l=1}^{n} b_{kl}a_{lk}\right)=\sum_{k=1}^{n} d_{kk}=\operatorname{tr}(BA)$.

Theorem

Let A and B be two similar matrices. Then:

1) $\det A=\det B$.
2) $\operatorname{tr} A=\operatorname{tr} B$.
3) $c_A(x)=c_B(x)$.
4) Eigenvalues of A = eigenvalues of B.

Proof:

A and B are similar, so there exists P an invertible matrix such that $A=P^{-1}BP$.

1) $\det A=\det(P^{-1}BP)=\det(P^{-1})\times\det(B)\times\det(P)=\dfrac{1}{\det P}\times\det B\times\det P=\det B$.
2) $\operatorname{tr} A=\operatorname{tr}(P^{-1}BP)=\operatorname{tr}((BP)(P^{-1}))=\operatorname{tr} B$.
3) $$\begin{aligned} c_A(x)&=\det(xI-A)=\det(xI-P^{-1}BP)=\det(xP^{-1}P-P^{-1}BP)=\det(P^{-1}(xI-B)P)\\ &=\det(P^{-1})\times\det(xI-B)\times\det(P)=c_B(x)\end{aligned}.$$
4) We know $c_A(x)=c_B(x)$. Since the eigenvalues of A are the roots of $c_A(x)$, and the eigenvalues of B are the roots of $c_B(x)$, we have eigenvalues of A = eigenvalues of B.

DIAGONALIZATION

Theorem

1) A is diagonalizable iff there exists a basis of eigenvectors of A. It means there exists eigenvectors $X_1, \dots, X_n$ such that $P = (X_1 \quad \cdots \quad X_n)$ is invertible.
2) In this case, $D = P^{-1}AP = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ where λ_j is the eigenvalue corresponding to X_j.

Method For Diagonalizing a Matrix

1) Find the eigenvalues – compute $c_A(x)$ and find the roots.
2) If λ is an eigenvalue, then find the basic solution of $(\lambda I - A)X = 0$.
3) Do we have a basis of eigenvectors?
   - If no, then A is not diagonalizable.
   - If yes, then $X_1, \dots, X_n$ is a basis of eigenvectors. $P = (X_1 \quad \cdots \quad X_n)$, and $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ such that $AX_j = \lambda_j X_j$.

Example

Diagonalize $A = \begin{pmatrix} 5 & -4 \\ 6 & -5 \end{pmatrix}$.

- $c_A(x) = \begin{vmatrix} x-5 & 4 \\ -6 & x+5 \end{vmatrix} = x^2 - 1$. So $\lambda_1 = -1$, $\lambda_2 = 1$. Since $\lambda_1 \neq \lambda_2$, the eigenvectors X_1, X_2 are linearly independent.
- -1-eigenvectors: $(-I - A)X = 0 \Leftrightarrow \begin{cases} -6x_1 + 4x_2 = 0 \\ -6x_1 + 4x_2 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = \frac{2}{3}t \\ x_2 = t \end{cases}, t \in \mathbf{R}$. So $X_1 = \begin{pmatrix} \frac{2}{3} \\ 1 \end{pmatrix}$.
- 1-eigenvectors: $(I - A)X = 0 \Leftrightarrow \begin{cases} -4x_1 + 4x_2 = 0 \\ -4x_1 + 4x_2 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = t \\ x_2 = t \end{cases}, t \in \mathbf{R}$. So $X_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
- $P = (X_1 \quad X_2) = \begin{pmatrix} \frac{2}{3} & 1 \\ 1 & 1 \end{pmatrix}$. So we have $P^{-1}AP = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$.

Example

Diagonalize $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

- $c_A(x) = \begin{vmatrix} x-1 & -2 & 0 \\ 0 & x-1 & 0 \\ 0 & 0 & x-3 \end{vmatrix} = (-1)^{3+3}(x-3)(x-1)^2 = (x-3)(x-1)^2$. So $\lambda_1 = 3$, $\lambda_2 = 1$. (<u>Remark</u>: If A is a triangle matrix, then the eigenvalues of A are the entries on the main diagonal)
- 3-eigenvectors: $(3I - A)X = 0 \Rightarrow X = s\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, s \in \mathbf{R}$. So $X_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

- 1-eigenvectors: $(I-A)X=0 \Rightarrow X=s\begin{pmatrix}1\\0\\0\end{pmatrix}, s\in \mathbf{R}$. So $X_2=\begin{pmatrix}1\\0\\0\end{pmatrix}$.
- Since there only X_1 and X_2 independent eigenvectors, it is not a basis of $\mathbf{R}^3$. So A is <u>not</u> diagonalizable.

MULTIPLICITY OF EIGENVALUES

Definition

Let A be an $n\times n$ matrix. Let λ be an eigenvector of A. Suppose $c_A(x)=0$. We can write $c_A(x)=(x-\lambda)^m \cdot g(x), g(x)\neq 0$. m is the <u>multiplicity</u> of λ.

Proposition

Let λ be an eigenvalue of A. Let m be its multiplicity. Then $\dim E_\lambda(A)\leq m$.

Lecture #22 – Thursday, March 25, 2004

Theorem

A is diagonalizable iff $\dim E_\lambda(A)=m(\lambda)$ (the multiplicity of λ) for all λ eigenvalue of A.

Proof:

- Assume A is diagonalizable. Prove $\dim E_\lambda(A)=m(\lambda)$.
  - $D=P^{-1}AP=\operatorname{diag}(\lambda_1,\ldots,\lambda_n)$, and $A=PDP^{-1}$.
  - $c_A(x)=\det(xI-A)=c_D(x)=\det(xI-D)$.
  - $c_A(x)=\begin{vmatrix} x-x_1 & \cdots & 0\\ \vdots & \ddots & \vdots \\ 0 & \cdots & x-x_n\end{vmatrix}=(x-x_1)(x-x_2)\cdots(x-x_n)$.
  - Let λ be an eigenvalue of A. The multiplicity of λ is $m(\lambda)=\#\{j, \lambda_j=\lambda\}$.
  - On the other hand, we know $P=\begin{pmatrix}X_1 & \cdots & X_n\end{pmatrix}$ where $AX_j=\lambda_j X_j$.
  - $E_\lambda(A)=\operatorname{span}\{X_j, AX_j=\lambda X_j\}$. So $\dim E_\lambda(A)=\#\{j, AX_j=\lambda X_j\}=\#\{j,\lambda_j=\lambda\}=m(\lambda)$.
- Assume $\dim E_\lambda(A)=m(\lambda)$. Prove A is diagonalizable.
  - Let $\lambda_1,\ldots,\lambda_k$ be the eigenvectors of A. Let $\alpha_1,\ldots,\alpha_k$ be the multiplicity of $\lambda_1,\ldots,\lambda_k$ such that $c_A(x)=(x-\lambda_1)^{\alpha_1}(x-\lambda_2)^{\alpha_2}\cdots(x-\lambda_k)^{\alpha_k}$.
  - By assumption, $\dim E_{\lambda_j}(A)=m(\lambda_j)=\alpha_j$.
  - Let $X_{\alpha_1}(\lambda_k),\ldots,X_{\alpha_k}(\lambda_k)$ be a basis of $E_{\lambda_k}(A)$.
  - $\{X_j(\lambda_l), i\leq j\leq \alpha_l, l=1,\ldots,k\}$ is a linearly independent set.
  - There are $\alpha_1+\cdots+\alpha_k$ vectors. So $\{X_j(\lambda_l)\}$ is a basis of $\mathbf{R}^n$. So A is diagonalizable.

Example

Show that $A=\begin{pmatrix}2&0&-1\\1&1&-1\\0&0&1\end{pmatrix}$ is diagonalizable.

- Compute the eigenvalues: $c_A(x)=\det(xI-A)=(x-1)^2(x-2)$. So $\lambda_1=1, m(\lambda_1)=2$, and $\lambda_2=2, m(\lambda_2)=1$.
- Eigenspace for λ_1: $E_{\lambda_1}(A)=\operatorname{span}\left\{\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix}\right\}$.
- Since $m(\lambda_1)=\dim E_{\lambda_1}(A)$, so A is diagonalizable.

PROPERTIES OF DIAGONALIZABLE MATRICES

Let A be a diagonalizable matrix.

1) If the eigenvalues of A are ±1, the A is invertible and $A^{-1}=A$.
   - Proof: There exists P such that $P^{-1}AP=\operatorname{diag}(\lambda_1,\ldots,\lambda_n), \lambda_j=\pm 1$. So $A=P\operatorname{diag}(\lambda_1,\ldots,\lambda_n)P^{-1}$. So $A^2=P(\operatorname{diag}(\lambda_1,\ldots,\lambda_n))^2P^{-1}=P\operatorname{diag}(\lambda_1{}^2,\ldots,\lambda_n{}^2)P^{-1}$. Since $\lambda_j=\pm 1$, $\lambda_j{}^2=1$. So $A^2=P\operatorname{diag}(1,\ldots,1)P^{-1}=PIP^{-1}=I$. So $A^{-1}=A$.

2) If the eigenvalues of A are 0 or 1, then $A^2=A$.
   - Proof: $A=P\operatorname{diag}(\lambda_1,\ldots,\lambda_n)P^{-1}, \lambda_j=0,1$. So $A^2=P\operatorname{diag}(\lambda_1{}^2,\ldots,\lambda_n{}^2)P^{-1}$. Since $\lambda_j=0$ or $\lambda_j=1$, $\lambda_j{}^2=\lambda_j$. So $A^2=P\operatorname{diag}(\lambda_1,\ldots,\lambda_n)P^{-1}=A$.

3) If the eigenvalues of A are non-negative, then there exists a matrix B such that $A=B^2$.
   - Proof: $A=P\operatorname{diag}(\lambda_1,\ldots,\lambda_n)P^{-1}, \lambda_j\ge 0$. Define $\mu_j=\sqrt{\lambda_j}\,\forall j$. Let $B=P\operatorname{diag}(\mu_1,\ldots,\mu_n)P^{-1}\Rightarrow B^2=P\operatorname{diag}(\mu_1{}^2,\ldots,\mu_n{}^2)P^{-1}=P\operatorname{diag}(\lambda_1,\ldots,\lambda_n)P^{-1}=A$.
   - Note: Can also take $\pm\mu_j$.

4) If $A^k=0$ for some k, then $A=0$.
   - Proof: $A=P\operatorname{diag}(\lambda_1,\ldots,\lambda_n)P^{-1}$. $A^k=P\operatorname{diag}(\lambda_1{}^k,\ldots,\lambda_n{}^k)P^{-1}=0\Rightarrow\lambda_1{}^k=\cdots=\lambda_n{}^k=0\Rightarrow\lambda_1=\cdots=\lambda_n=0$. So $A=P0P^{-1}=0$.
   - Remark: If $A^k=0$ for a matrix A non-diagonalizable, then the eigenvalues of A are 0, so $c_A(x)=x^n$.

Proposition

Let A be a diagonalizable matrix. Let $\lambda_1,\ldots,\lambda_n$ be its eigenvalues. Then

- $\det A=\lambda_1\times\lambda_2\times\cdots\times\lambda_n$
- $\operatorname{tr} A=\lambda_1+\lambda_2+\cdots+\lambda_n$.

Proof:

- A is diagonalizable means $A = P\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)P^{-1}$ for P invertible.
- $\det A = \det\left(P\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)P^{-1}\right) = \det P \times \det(\mathrm{diag}(\lambda_1, \ldots, \lambda_n)) \times \det P^{-1} = \lambda_1 \times \cdots \times \lambda_n$.
- $\mathrm{tr}\, A = \mathrm{tr}\left(P\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)P^{-1}\right) = \mathrm{tr}\left(\left(\mathrm{diag}(\lambda_1, \ldots, \lambda_n)P^{-1}\right)P\right) == \mathrm{tr}(\mathrm{diag}(\lambda_1, \ldots, \lambda_n)) = \lambda_1 + \cdots + \lambda_n$.

Lecture #23 – Tuesday, March 30, 2004

FUNCTION SPACES

Definition

A set V of object is a vector space if:

1) There exists an operation $\oplus$ on $V \times V$ such that:
   - (A1) If u and v are in V, then $u \oplus v$ is in V.
   - (A2) $u \oplus v = v \oplus u$ for all u and v are in V (commutative).
   - (A3) $u \oplus (v \oplus w) = (u \oplus v) \oplus w$ for all u, v, w in V (associative).
   - (A4) An element 0 in V exists such that $v \oplus 0 = v = 0 \oplus v$ for every v in V.
   - (A5) For each v in V, an element $-v$ in V exists such that $-v \oplus v = 0 = v \oplus -v$.
2) There exits a multiplication by scalar $\otimes$ such that:
   - (S1) For $a \in \mathbf{R}$ and $v \in V$, $a \otimes v \in V$.
   - (S2) $a \otimes (u \oplus v) = a \otimes u \oplus a \otimes v$ for $a \in \mathbf{R}$ and $u, v \in V$.
   - (S3) $(a + b) \otimes v = a \otimes v \oplus b \otimes v$ for $a, b \in \mathbf{R}$ and $v \in V$.
   - (S4) $(ab) \otimes v = a \otimes (b \otimes v)$ for $a, b \in \mathbf{R}$ and $v \in V$.
   - (S5) $1 \otimes v = v$ for $v \in V$.

Example

$V = \mathbf{R}^n$ is a vector space if we define $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \oplus \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}$, $a \otimes \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}$, $0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$, $-\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} -x_1 \\ \vdots \\ -x_n \end{pmatrix}$.

Non-Example

Let $V = \mathbf{R}^2$. Define $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_2 + y_2 \\ x_1 + y_1 \end{pmatrix}$ and $a \otimes \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} ax_1 \\ ax_2 \end{pmatrix}$. Is V still a vector space?

- (A1): Let $u, v \in \mathbf{R}^2$. It is clear that $u \oplus v \in \mathbf{R}^2$.
- (A2): Let $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $u = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$. $u \oplus v = \begin{pmatrix} u_2 + v_2 \\ u_1 + v_1 \end{pmatrix} = \begin{pmatrix} v_2 + u_2 \\ v_1 + u_1 \end{pmatrix} = v \oplus u$.

- (A3): Let $w=\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$. $(u\oplus v)\oplus w=\begin{pmatrix} u_2+v_2 \\ u_1+v_1 \end{pmatrix}\oplus\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}=\begin{pmatrix} u_1+v_1+w_2 \\ u_2+v_2+w_1 \end{pmatrix}$, but $u\oplus(v\oplus w)=\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}\oplus\begin{pmatrix} v_2+w_2 \\ v_1+w_1 \end{pmatrix}=\begin{pmatrix} u_2+v_2+w_1 \\ u_1+v_1+w_2 \end{pmatrix}$, so $u\oplus(v\oplus w)\neq(u\oplus v)\oplus w$.
- So $(V,\oplus,\otimes)$ is not a vector space.

Example

Consider $F[a,b]:\{f:[a,b]\to\mathbf{R}\}$. Define $(f\oplus g)(x)=f(x)+g(x)$ and $(a\otimes g)(x)=a\times f(x)$. Is $F[a,b]$ a vector space?

- (A1): Let *f*, $g\in$ $F[a,b]$. It is clear that $f\oplus g\in F[a,b]$.
- (A2): $(f\oplus g)(x)=f(x)+g(x)=g(x)+f(x)=(g\oplus f)(x)$. So $f\oplus g=g\oplus f$.
- (A3): Let *f*, *g*, $h\in$ $F[a,b]$. $\begin{aligned}&(f\oplus(g\oplus h))(x)=f(x)+(g\oplus h)(x)=f(x)+(g(x)+h(x))\\&=(f(x)+g(x))+h(x)=(f\oplus g)(x)+h(x)=((f\oplus g)\oplus h)(x)\end{aligned}$. So $f\oplus(g\oplus h)=(f\oplus g)\oplus h$.
- (A4): Let $(f\oplus g_0)(x)=f(x)\Leftrightarrow f(x)+g_0(x)=f(x)\Rightarrow g_0(x)=0$. So $g_0(x)=0(x)$ for $x\in[a,b]$.
- (A5): Let $(f\oplus g)(x)=f(x)+g(x)=0(x)=0\Leftrightarrow g(x)=-f(x)$. So *-f* is defined by $(-f)(x)=-f(x)$.
- Similarly, $F[a,b]$ will satisfy (S1) to (S5). So $F[a,b]$ is a vector space.

Example

Let *P* be the set of polynomials. $P=\{p(x)=a_0+a_1x+\cdots+a_nx^n, a_j\in\mathbf{R}, n\in\mathbf{N}\}$. If we define $(p\oplus q)(x)=p(x)+q(x)$ and $(a\otimes p)(x)=a\times p(x)$, then it can be proven that $(P,\oplus,\otimes)$ is a vector space.

SUBSPACES

Definition

Let *V* be a vector space. Let $U\subset V, U\neq\{\ \}$. Then *U* is a <u>subspace</u> if:

(P1): $u\oplus v\in U$ for all *u*, $v\in$ *V* (stability of addition).
(P2): $a\otimes v\in U$ for all $a\in$ **R** and $v\in$ *V* (stability of addition).

Remark

1) *U* is a vector space.
2) $\{0\}\in U$.

Example

Consider $U\subset P$, a set of polynomials with root 4. $U=\{p\in P, p(4)=0\}$. Show that *U* is a subspace.

- $0\in U$ because $p(0)=0$.
- (P1): Let *p*, $q\in$ *U*. It means $p(4)=0$ and $q(4)=0$. $(p\oplus q)(4)=p(4)+q(4)=0$, so $p\oplus q\in U$.
- (P2): Let $a\in$ **R** and $p\in$ *U*. $(a\otimes p)(4)=a\times p(4)=a\times 0=0$, so $a\otimes p\in U$.

Example

The set of polynomials of degree less than n is a subspace of P.

Example

Let $Q_n \subset P$ be the set of polynomial of degree n. Show that Q_n is not a subspace.

- Define $p_1 = x^n \in Q_n$ and $p_2 = -x^n \in Q_n$.
- $p_1 \oplus p_2 = 0 \notin Q_n$ for $n \geq 1$.
- So Q_n is not a subspace when $n \geq 1$. For $n = 0$, Q_n is a subspace.

Example

Let $U \subset F[a,b]$ be the set of differentiable functions. Is U a subspace?

- (P1): Let $f, g \in U$. $(f \oplus g)'(x) = f'(x) + g'(x)$, so $f \oplus g \in U$.
- (P2): Let $a \in \mathbf{R}$ and $f \in U$. $(a \otimes f)'(x) = a \times f'(x)$, so $a \otimes f \in U$.
- Therefore, U is a subspace.

Examples

1) Let $U \subset F[0,1]$ such that $f(0) = f(1)$. Is U a subspace?
   - $0(0) = 0 = 0(1)$. So $0 \in U$.
   - (P1): Let $f, g \in U$. It means $f(0) = f(1)$ and $g(0) = g(1)$.
     $(f+g)(0) = f(0) + g(0) = f(1) + g(1) = (f+g)(1)$. So $f + g \in U$.
   - (P2): Let $a \in \mathbf{R}$ and $f \in U$. $(a \times f)(0) = a \times f(0) = a \times f(1) = (a \times f)(1)$. So $a \times f \in U$.
   - Therefore, U is a subspace.

2) Let $U \subset F[0,1]$ such that $f(0) = 1$. Is U a subspace?
   - $0(1) = 0 \neq 1$, so $0 \notin U$. Therefore, U is not a subspace.

INDEPENDENT SETS OF VECTOR SPACE

Definition

Let $\{v_1, v_2, \ldots, v_n\}$ be a set in a vector space V. This set is (linearly)independent if
$s_1 v_1 + s_2 v_2 + \cdots + s_n v_n = 0 \Rightarrow s_1 = s_2 = \cdots = s_n = 0$.

Proposition

1) If $v \in V, v \neq 0$, then $\{v\}$ is independent.
2) If $\{v_1, \ldots, v_n\}$ an independent set, then $\{a_1 v_1, \ldots, a_n v_n\}, a_j \neq 0$ is also independent.
3) If $\{v_1, \ldots, v_n\}$ is independent, then $v_j \neq 0$.

Proof of (1):

- If $sv = 0$ and $s \neq 0$, then $\frac{1}{s} sv = \frac{1}{s} 0 \Rightarrow 1 \cdot v = v = 0$. But since $v \neq 0$, this is impossible. So $s = 0$ and $\{v\}$ is independent.

Proof of (2):

- Let $s_1(a_1v_1)+\cdots+s_n(a_nv_1)=0 \overset{\text{S4}}{\Leftrightarrow} (s_1a_1)v_1+\cdots+(s_na_n)v_n=0$. Since $\{v_1,\ldots,v_n\}$ is independent, we have $s_1a_1=\cdots=s_na_n=0$. Since $a_j\neq 0$, $s_ja_j=0\Rightarrow s_j=0$. So $\{a_1v_1,\ldots,a_nv_n\}, a_j\neq 0$ is independent.

Proof of (3):

- Suppose $v_k=0$. Then $s_1v_1+\cdots+s_nv_n=0$ for $s_j=0, j=k$ and $s_k=1$, so $\{v_1,\ldots,v_n\}$ are not independent. Therefore, $v_j\neq 0, j=1,\ldots,n$.

Example

Let $P=\{1,x,x^2,\ldots,x^n\}$. Is P independent?

- Let $p(x)=s_0\cdot 1+s_1x+s_2x^2+\cdots+s_nx^n=0$. Since $p(x)$ has an infinite number of roots, $p(x)$ has to be the polynomial 0. So $s_0=\cdots=s_n=0$.

Example

Let $\{p_1,p_2,\ldots,p_n\}$ be polynomials of distinct degree. Show that this set is independent.

- Let d_j be the degree of p_j. Assume that $d_1<d_2<\cdots<d_n$.
- Let $Q=s_1p_1+s_2p_2+\cdots+s_np_n=0$. Let ax^{d_n} be the term of degree d_n of p_n, so $a\neq 0$.
- Then the term of degree d_n of Q is $s_n\cdot a\cdot x^{d_n}$, so $s_n\cdot a=0\Rightarrow s_n=0$.
- So $Q=s_1p_1+s_2p_2+\cdots+s_{n-1}p_{n-1}=0$. Apply the same idea for s_{n-1}.

Example

Consider $\operatorname{im} P_2=\{x^2-1,x-1,x+2\}$. Is $\operatorname{im} P_2$ independent?

- Let $s_1(x^2-1)+s_2(x-1)+s_3(x+2)=0 \Leftrightarrow s_1x^2+(s_2+s_3)x+(-s_1-s_2+2s_3)=0$, so
$$\begin{cases} s_1=0 \\ s_2+s_3=0 \\ -s_1-s_2+2s_3=0 \end{cases} \Leftrightarrow \begin{cases} s_1=0 \\ s_2=0 \\ s_3=0 \end{cases}.$$
- So $\operatorname{im} P_2=\{x^2-1,x-1,x+2\}$ is independent.

Example

Is $\operatorname{im} F[0,2\pi]=\{\cos x,\sin x\}$ independent?

- Let $s_0\cos x+s_1\sin x=0, x\in[0,2\pi]$.
- When $x=0$, $s_0\cos 0+s_1\sin 0=0\Rightarrow s_0=0$.
- When $x=\frac{\pi}{2}$, $s_0\cos\frac{\pi}{2}+s_1\sin\frac{\pi}{2}=0\Rightarrow s_1=0$.
- So $\operatorname{im} F[0,2\pi]=\{\cos x,\sin x\}$ is independent.

Spanning Sets of Vector Space

Definition

Let $\{v_1,v_2,\ldots,v_n\}$ be in the vector space *V*. $U=\operatorname{span}\{v_1,\ldots,v_n\}$ is the set of all linear combinations of the v_j's. We say that *U* is <u>spanned</u> by the v_j's.

Proposition

$U = \text{span}\{v_1, \ldots, v_n\}$ is a subspace of *V*.

Example

Let $U = \left\{ \begin{pmatrix} a+b & b \\ -b & a \end{pmatrix}, a, b \in \mathbf{R} \right\} \subset M_{2\times2} = \{2\times 2 \text{ matrices}\}$. Show that *U* is a spanning set.

- If $A \in U$, $A = \begin{pmatrix} a+b & b \\ -b & a \end{pmatrix}$ for *a*, *b* ∈ **R**. So $A = a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$.
- So $A \in \text{span}\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \right\}$, and $U \subset \text{span}\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \right\}$.
- For $a = 1$ and $b = 0$, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in U$. For $a = 0$ and $b = 1$, $\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \in U$.
- So $U = \text{span}\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} \right\}$.

Lecture #24 – Thursday, April 1, 2004

Basis of Vector Spaces

Definition

$\{e_1, e_2, \ldots, e_n\} \subset V$ is a <u>basis</u> of *V* if:

(B1) $\{e_1, \ldots, e_n\}$ are independent.

(B2) $V \subset \text{span}\{e_1, \ldots, e_n\}$.

Theorem

If $\{e_1, \ldots, e_n\}$ is a basis of *V*, then any other basis has *n* vectors. We say that $\dim V = n$.

Example

Let $M_{mn} = \{m \times n \text{ matrices}\}$. Show $\dim M_{mn} = m \times n$.

- Let E_{kl} be matrices where $E_{kl} = \left(e_{ij}^{kl}\right) = \begin{pmatrix} e_{11}^{kl} & \cdots & e_{1n}^{kl} \\ \vdots & \ddots & \vdots \\ e_{n1}^{kl} & \cdots & e_{nn}^{kl} \end{pmatrix}$ such that $e_{ij}^{kl} = 0$ for $i \neq k$, $j \neq l$, and $e_{ij}^{kl} = 1$.
- (B1): Want $\{E_{kl}\}, \begin{matrix} 1 \le k \le m \\ 1 \le l \le n \end{matrix}$ is independent. We write $\sum_{k,l} \alpha_{kl} E_{kl} = 0, \alpha_{kl} \in \mathbf{R}$. We can compute $\sum_{k,l} \alpha_{kl} E_{kl} = (\alpha_{kl}) = 0 \Rightarrow \alpha_{kl} = 0 \forall k, l$. So $\{E_{kl}\}$ are independent.

- (B2): Let $A \in M_{mn}$ and a_j be the (*i*, *j*)-entry of *A*, so $A = (a_{ij}) \Rightarrow A = \sum_{i,j} a_{ij} E_{ij}$. So $\{E_{kl}\}$ span M_{mn}.
- Since $\#\{E_{kl}, 1 \le k \le m, 1 \le l \le n\} = m \times n$, so $\dim M_{mn} = m \times n$.

Example

Let $P_n \subset P, P_n = \{a_0 + a_1 x + \cdots + a_n x^n\}, a_j \in \mathbf{R}$. Prove $\{1, x, \ldots, x^n\}$ is a basis of P_n.

- $\dim P_n = n+1$.
- (B1): $\{1, x, \ldots, x^n\}$ is linearly independent (a polynomial with distinct degrees).
- (B2): $P_n = \{a_0 + a_1 x + \cdots + a_n x^n\} = \operatorname{span}\{1, x, \ldots, x^n\}$.
- So $\{1, x, \ldots, x^n\}$ is the standard basis of P_n. Also, $\#\{1, x, \ldots, x^n\} = n+1$, so $\dim P_n = n+1$.

Example

Let $U \subset M_{22}, U = \{A \in M_{22}, A^T = -A\}$. Find $\dim U$.

- Let $A \in U$. So $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and $A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, $-A = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$.

  $A^T = -A \Leftrightarrow \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix} \Leftrightarrow \begin{cases} a = -a \\ c = -b \\ b = -c \\ d = -d \end{cases} \Leftrightarrow \begin{cases} a = 0 \\ c = -b \\ d = 0 \end{cases}$. So $A = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} = b\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = bM_0$.

  So $U \subset \operatorname{span}\{M_0\}$.
- Let $B \in \operatorname{span}\{M_0\}$. So $B = bM_0 = b\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} \Rightarrow B^T = -B$, so $B \in U$. So $\operatorname{span}\{M_0\} \subset U$.
- $bM_0 = 0 \Rightarrow b = 0$, so M_0 is independent (B1). Also $\operatorname{span}\{M_0\} = U$ (B2). So M_0 is a basis of *U*. So $\dim U = 1$.

Example

Consider $P_3 \subset P, P_3 = \{a_0 + a_1 x + a_2 x^2 + a_3 x^3\}, a_j \in \mathbf{R}$. Find a basis of P_3 that contains $\{1 + x, 2 + x^2\}$.

- Since $\dim P_3 = 4$, so we need 2 more vectors.
- $\{1, x, x^2, x^3\}$ is a basis of P_3.
- Pick 1 and x^3. $\{1, 1+x, 2+x^2, x^3\}$ are independent because they are of different degrees.
- Since $\{1, 1+x, 2+x^2, x^3\}$ are independent and $\#\{1, 1+x, 2+x^2, x^3\} = 4 = \dim P_3$, there is no need to look for $\operatorname{span}\{1, 1+x, 2+x^2, x^3\}$. So $\{1, 1+x, 2+x^2, x^3\}$ is a basis of P_3.

Lecture #25 – Tuesday, April 6, 2004

Proposition

If *f* and *g* are polynomials such that $f \cdot g = 0$, then either $f = 0$ or $g = 0$.

Proof:

- Suppose $f \neq 0$ and $g \neq 0$. Denote $x_1, \dots, x_n$ the roots of *f*, and $y_1, \dots, y_k$ the roots of *g*.
- Let $x \notin \{x_j, j = 1, \dots, n, y_l, l = 1, \dots, k\}$. Then $\begin{cases} f(x) \neq 0 \\ g(x) \neq 0 \end{cases}$, so $(f \cdot g)(x) \neq 0$. Contradiction!

Example

Let $U = \{(x^2 - x)p(x), p \in P_2\} \subset P_4$. Find a basis of *U*.

- Let $q \in U$. So
$$\begin{aligned} q(x) &= (x^2 - x) \cdot p(x), p \in P_2 \\ &= (x^2 - x)(a_0 + a_1 x + a_2 x^2), a_0, a_1, a_2 \in \mathbf{R}. \\ &= a_0(x^2 - x) + a_1 x(x^2 - x) + a_2 x^2(x^2 - x) \end{aligned}$$
- So $q \in \operatorname{span}\{(x^2 - x), x(x^2 - x), x^2(x^2 - x)\} := \operatorname{span}\{p_0, p_1, p_2\}$. Since $p_0, p_1, p_2 \in U$, so $\operatorname{span}\{p_0, p_1, p_2\} \subset U$.
- Since the degrees of p_0, p_1, p_2 are of distinct degrees, so p_0, p_1, p_2 are independent.
- So a basis of *U* is $\{p_0, p_1, p_2\}$ and $\dim(U) = 3$.

Example

Let $U = \{p \in P_n, p(a) = 0\}, a \in \mathbf{R}$. Show $\{(x-a), (x-a^2), \dots, (x-a)^n\}$ is a basis of *U*.

- Since $\{(x-a), (x-a^2), \dots, (x-a)^n\}$ have distinct degrees, they are independent.
- Let $p_k(x) = (x-a)^k, k = 1, \dots, n$. $p_k(a) = 0$, so $p_k \in P_n, \forall k \leq n$. So $p_k \in U$.
- So $\operatorname{span}\{p_1, p_2, \dots, p_n\} \subset U$. So $\dim U \geq n$.
- Also, $U \subset P_n \Rightarrow \dim U \leq \dim P_n = n + 1$.
- If $\dim U = n + 1$, it means that $U \subset P_n$. But $\dim U = \dim P_n$ means $U = P_n$. This is impossible. Consider $p(x) = 1 \in P_n$, but $p \notin U$ since $p(0) \neq 0$.
- So the only possibility is $\dim U = n$. So $\{p_1, p_2, \dots, p_n\} = \{(x-a), (x-a^2), \dots, (x-a)^n\}$ is a basis of *U*.

VECTOR SPACES OF SOLUTIONS

Definition

Let *f* be a function of a variable $x \in \mathbf{R}$. Let $f^{(n)}$ be the n^{th} derivative of *f*. An equation of the form $f^{(n)} + a_{n-1} f^{(n-1)} + a_{n-2} f^{(n-2)} + \cdots + a_1 f' + a_0 f = 0, a_0, \dots, a_n \in \mathbf{R}$ is called <u>differential equation of order *n*</u>.

Definition

Finding the solution of a differential equation $f^{(n)} + a_{n-1} f^{(n-1)} + \cdots + a_1 f' + a_0 f = 0$ such that

$$(I) \begin{cases} f(x_0) = f \\ f'(x_0) = f_1 \\ \vdots \\ f^{(n-1)}(x_0) = f_{n-1} \end{cases}$$

is called an <u>initial value problem</u>. The conditions (I) are called the <u>initial condition</u>.

FIRST ORDER DIFFERENTIAL EQUATIONS

Equations of the form $f' + af = 0$ for $a \in \mathbf{R}$.

Theorem

Let U be the set of solution of $f' + af = 0$. Then:

1) U is a subspace of $F(\mathbf{R}) = \{f : \mathbf{R} \to \mathbf{R}\}$.
2) $\{e^{-ax}\}$ is a basis of U. $\dim U = 1$.

Proof (1):

- Let $U = \{f, f' + af = 0\}$. Let $f, g \in U$, so $f' + af = 0$ and $g' + ag = 0$.
- $(f+g)' + a(f+g) = f' + g' + af + ag = f' + af + g' + ag = 0 + 0 = 0$. So $f + g \in U$.
- Let $\alpha \in \mathbf{R}$. $(\alpha f)' + a(\alpha f) = \alpha f' + a\alpha f = \alpha(f' + af) = \alpha(0) = 0$. So $\alpha f \in U$.
- So U is a subspace.

Proof (2):

- Let $f_a(x) = e^{-ax}$.
- $f_a'(x) = -ae^{-ax}$. So $f_a'(x) + af_a(x) = -ae^{-ax} + ae^{-ax} = 0$. So $f_a \in U \Rightarrow \text{span}(f_a) \subset U$.
- Let $g \in U$. So $g' + ag = 0$. Write $g(x) = e^{-ax} \cdot h(x) \Rightarrow g'(x) = -ae^{-ax} \cdot h(x) + e^{-ax} \cdot h'(x)$. So $0 = g'(x) + ag(x) = -ae^{-ax} \cdot h(x) + e^{-ax} \cdot h'(x) + ae^{-ax} \cdot h(x) = e^{-ax} \cdot h'(x) \Leftrightarrow h'(x) = 0 \Leftrightarrow h(x) = c$. Finally, $g(x) = e^{-ax} \cdot h(x) = ce^{-ax}$, so $g \in \text{span}(e^{-ax}) \Rightarrow U \subset \text{span}(f_a)$.
- Therefore, $U = \text{span}(f_a)$.
- Since f_a is also independent, so $\{f_a\} = \{e^{-ax}\}$ is a basis of U, and $\dim U = 1$.

Examples

Solve the initial value problems.

1) $f' + af = 0$ and $f(0) = c_0$.
   - Let f_1 be the solution. So $f_1(x) = c \times e^{-ax}$.
   - $f_1(0) = c \times e^{-a0} = c$, so $c = c_0$.
   - So $f_1(x) = c_0 e^{-ax}$.

2) $f' + af = 0$ and $f(2) = c_0$.
   - Let f_2 be the solution. So $f_2(x) = c \times e^{-ax}$.
   - $f_2(2) = c \times e^{-2a} \Rightarrow c_0 = ce^{-2a} \Rightarrow c = c_0 e^{2a}$.
   - So $f_2(x) = c_0 e^{2a} \times e^{-ax} = c_0 e^{-a(x-2))}$.

SECOND ORDER DIFFERENTIAL EQUATIONS

Equations of the form $f'' + a_1 f' + a_2 f = 0$.

Definition

The <u>characteristic polynomial</u> of this equation is $p(x) = x^2 + a_1 x + a_2$.

Theorem

Let U be the set of solutions of $f'' + a_1 f' + a_2 f = 0$. Let λ and μ be the roots of $p(x) = x^2 + a_1 x + a_2$. Then:

1) If $\lambda \neq \mu$, then $\{e^{\lambda x}, e^{\mu x}\}$ is a basis of U.
2) If $\lambda = \mu$, then $\{e^{\lambda x}, \lambda e^{\lambda x}\}$ is a basis of U.

Proof of (1): $\lambda \neq \mu$

- We know U is a subspace. We admit $\dim U = 2$.
- Let $f_\lambda(x) = e^{\lambda x}$. So $f'_\lambda(x) = \lambda e^{\lambda x}$, and $f''_\lambda(x) = \lambda^2 e^{\lambda x}$.
  $f''_\lambda(x) + a_1 f'_\lambda(x) + a_2 f_\lambda(x) = \lambda^2 e^{\lambda x} + \lambda e^{\lambda x} + e^{\lambda x} = e^{\lambda x}(\lambda^2 + a_1\lambda + a_2) = e^{\lambda x} p(\lambda) = 0$, so $f_\lambda \in U$.
  Similarly, $f_\mu(x) = e^{\mu x} \in U$.
- Let $s_1 f_\lambda(x) + s_2 f_\mu(x) = 0 \Rightarrow s_1 e^{\lambda x} + s_2 e^{\mu x} = 0$. Differentiating both sides, we obtain $s_1 \lambda e^{\lambda x} + s_2 \mu e^{\mu x} = 0$. Now let $x = 0$. Then $\begin{cases} s_1 + s_2 = 0 \\ \lambda s_1 + \mu s_2 = 0 \end{cases} \Leftrightarrow \begin{cases} s_2 = -s_1 \\ (\lambda - \mu)s_1 = 0 \end{cases} \Leftrightarrow \begin{cases} s_1 = 0 \\ s_2 = 0 \end{cases}$. So $\{f_\lambda, f_\mu\}$ are independent.
- Therefore, $\{f_\lambda(x), f_\mu(x)\} = \{e^{\lambda x}, e^{\mu x}\}$ is a basis of U.

Proof of (2): $\lambda = \mu$

- Let $f_\lambda(x) = e^{\lambda x}$ and $g_\lambda(x) = xe^{\lambda x}$. So $g'_\lambda(x) = e^{\lambda x} + x(\lambda e^{\lambda x}) = e^{\lambda x}(1 + \lambda x)$ and
  $g''_\lambda(x) = \lambda e^{\lambda x}(1 + \lambda x) + e^{\lambda x}(\lambda) = e^{\lambda x}(x\lambda^2 + 2\lambda)$.
- $$\begin{aligned} g''_\lambda(x) + a_1 g'_\lambda(x) + a_2 g(x) &= e^{\lambda x}(x\lambda^2 + 2\lambda) + a_1 e^{\lambda x}(1 + \lambda x) + a_2 x e^{\lambda x} \\ &= e^{\lambda x}\left(x(\lambda^2 + a_1\lambda + a_2) + 2\lambda + a_1\right) = e^{\lambda x}(x \cdot p(\lambda) + p'(\lambda)) = 0 \end{aligned}$$ So
  $g_\lambda \in U$.
- Let $s_1 f_\lambda(x) + s_2 g_\lambda(x) = 0 \Rightarrow s_1 e^{\lambda x} + s_2 x e^{\lambda x} = 0 \Rightarrow e^{\lambda x}(s_1 + x s_2) = 0 \Rightarrow s_1 + x s_2 = 0 \Rightarrow \begin{cases} s_1 = 0 \\ s_2 = 0 \end{cases}$.
  So $\{f_\lambda, g_\lambda\}$ are independent.
- So $\{f_\lambda(x), g_\lambda(x)\} = \{e^{\lambda x}, xe^{\lambda x}\}$ is a basis of U.

Example

Find all the solutions to $f'' - 4f' + 3f = 0$.

- The characteristic polynomial is $p(x) = x^2 - 4x + 3 = (x-3)(x-1)$. The roots are 1 and 3.
- So the solutions are given by $f(x) = c_1 e^x + c_2 e^{3x}, c_1, c_2 \in \mathbf{R}$.

Example

Find the solution to $f'' + 2f' + f = 0$ such that $f(1) = 1$, $f'(1) = 0$.

- The characteristic polynomial $p(x) = x^2 + 2x + 1 = (x+1)^2$. So the only root is -1.
- The solution $f(x) = c_1 e^{-x} + c_2 x e^{-x} = e^{-x}(c_1 + c_2 x), c_1, c_2 \in \mathbf{R}$, so
  $f(1) = 1 \Rightarrow e^{-1}(c_1 + c_2) = 1 \Rightarrow c_1 + c_2 = e$. Also,

$f'(x) = -e^{-x}(c_1 + c_2 x) + e^{-x}(c_2) = e^{-x}(-c_1 - c_2 x + c_2)$, so $f'(1) = 0 \Rightarrow e^{-1}(-c_1) = 0 \Rightarrow c_1 = 0$.

So $\begin{cases} c_1 = 0 \\ c_1 + c_2 = e \end{cases} \Rightarrow \begin{cases} c_1 = 0 \\ c_2 = e \end{cases}$.

- So the solution is $f(x) = exe^{-x} = xe^{1-x}$.

Theorem

Let U be the solutions of $f'' + a_1 f' + a_2 f = 0$. If $p(x) = x^2 + a_1 x + a_2$ has a complex root $\lambda = p + qi, p, q \in \mathbf{R}$ ($p - qi$ is the second root), then $\{e^{px}\cos(qx), e^{px}\sin(qx)\}$ is a basis of U.

Proof:

- Let $f_{pq}(x) = e^{px}\cos(qx)$. Then $f'_{pq}(x) = e^{px}(p\cos(qx) - q\sin(qx))$ and $f''_{pq}(x) = e^{px}((p^2 - q^2)\cos(qx) - 2pq\sin(qx))$.
- $f''_{pq}(x) + a_1 f'_{pq}(x) + a_2 f_{pq}(x)$
  $= e^{px}((p^2 - q^2)\cos(qx) - 2pq\sin(qx)) + a_1 e^{px}(p\cos(qx) - q\sin(qx)) + a_2 e^{px}\cos(qx)$. Since
  $= e^{px}(\cos(qx)(p^2 - q^2 + a_1 p + a_2) + \sin(qx)(-2pq - a_1 q)) = 0$
  $p(p + iq) = (p^2 - q^2 + a_1 p + a_2) + i(-pq + a_1 q)$, $f''_{pq}(p + qi) + a_1 f'_{pq}(p + qi) + a_2 f_{pq}(p + qi) = 0$.
  so $f_{pq} \in U$. Similarly, $g_{pq}(x) = e^{px}\sin(qx) \in U$.
- Let $s_1 f_{pq} + s_2 g_{pq} = 0 \Rightarrow \begin{cases} s_1 = 0 \\ s_2 = 0 \end{cases}$. So they are linearly independent.
- Therefore, $\{f_{pq}(x), g_{pq}(x)\} = \{e^{px}\cos(qx), e^{px}\sin(qx)\}$ is a basis of U.

Corollary

Let $q \neq 0$. The solution to $f'' + q^2 f = 0$ are $U = \operatorname{span}\{\cos(qx), \sin(qx)\}$.

Proof:

- $p(x) = x^2 + q^2 = 0 \Rightarrow x = \pm iq$.
- Apply the Theorem with $p = 0$.

Example

Find the solution of $f'' - 4f' + 5f = 0$ such that $\begin{cases} f(0) = 1 \\ f\left(\frac{\pi}{2}\right) = 2 \end{cases}$.

- $p(x) = x^2 - 4x + 5 \Rightarrow x = 2 \pm i$.
- So $f(x) = ce^{2x}\cos x + de^{2x}\sin x, c, d \in \mathbf{R}$.
- Since $f(0) = 1 \Rightarrow ce^0\cos 0 + de^0\sin 0 = 1 \Rightarrow c = 1$, and $f\left(\frac{\pi}{2}\right) = 2 \Rightarrow ce^{\pi}\cos\frac{\pi}{2} + de^{\pi}\sin\frac{\pi}{2} = 2 \Rightarrow d = 2e^{-\pi}$.
- So $f(x) = e^{2x}\cos x + 2e^{-\pi}e^{2x}\sin x = e^{2x}\cos x + 2e^{2x-\pi}\sin x$.

Remark: The condition $\begin{cases} f(0) = 1 \\ f\left(\frac{\pi}{2}\right) = 2 \end{cases}$ is called a boundary condition.

Lecture #26 – Thursday, April 8, 2004

SYSTEMS OF DIFFERENTIAL EQUATIONS

Definition

Let $f_1, f_2, \dots, f_n$ be functions. The system of the form $\begin{cases} f_1' = a_{11}f_1 + a_{12}f_2 + \cdots a_{1n}f_n \\ \vdots \\ f_n' = a_{n1}f_1 + a_{n2}f_2 + \cdots a_{nn}f_n \end{cases}$, $a_{ij} \in \mathbf{R}$ is called a <u>system of differential equations</u>.

We write $f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$, $f' = \begin{pmatrix} f_1' \\ \vdots \\ f_n' \end{pmatrix}$, $A = (a_{ij})$. Then the system can be written as $f' = Af$.

Theorem

Let W be the set of vectors of functions. Then $(W, +, \times)$ is a vector space. Moreover:

1) $(f+g)' = f' + g'$.
2) $(af)' = af', a \in \mathbf{R}$.

Theorem

Assume A is diagonalizable. Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of A, and $X_1, X_2, \dots, X_n$ be the corresponding eigenvectors. Let U be the set of solutions of $f' = Af$. Then:

1) U is a subspace of W.
2) A basis of U is $\{X_1 e^{\lambda_1 x}, X_2 e^{\lambda_2 x}, \dots, X_n e^{\lambda_n x}\}$.

Example

Find all the solutions of $f' = Af$, where $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & -2 \\ 0 & -2 & 3 \end{pmatrix}$.

- $c_A(x) = \det(xI_3 - A) = \det\begin{pmatrix} x-1 & 0 & 0 \\ 0 & x-3 & 2 \\ 0 & 2 & x-3 \end{pmatrix} = (x-1)\left((x-3)^2 - 4\right) = (x-1)^2(x-5)$. So the eigenvalues are 5 and 1.
- Let $(5I_3 - A)X = 0 \Leftrightarrow \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \Leftrightarrow \begin{cases} 4x = 0 \\ 2y + 2z = 0 \\ 2y + 2z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 0 \\ y = -s, s \in \mathbf{R} \\ z = s \end{cases} \Leftrightarrow X = s\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$. So a 5-eigenvector is $X_5 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$.

- Similarly, the 1-eigenvectors are $X_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $Y_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.
- So the solutions are given by $f(x) = aX_5 e^{5x} + bX_1 e^{x} + cY_1 e^{x}, a, b, c \in \mathbf{R}$.