1201 Algebra 1 Notes

Based on the 2011 autumn lectures by Prof F E A Johnson

SOLVING EQUATIONS.

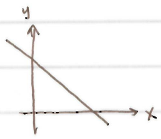
$2x+3=0$ linear in one variable

$3x^2+2x+3=0$ quadratic in one variable

$\left.\begin{array}{r} 3x^2+xy+y^2=5 \\ xy=1 \end{array}\right\}$ quadratic in two variables

LINEAR ALGEBRA

$x+y=1$ linear in two variables



geometric interpretation — represents a straight line in the x-y plane

$x+y+z=1$ linear in three variables.

imagine a particle in 3-space — 6 variables are needed to describe its motion

3 of position (x, y, z) 3 of momentum (a, b, c)

we represent variables by using single variable letters with subscripts for coordinates.

e.g. $\underset{\sim}{x} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7)$ ⟶ represents 7 independent coordinates. $\underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix}$

row vector column vector

how to write a single linear equation in 7 unknowns (variables)

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 = \underline{b}$$

$\underset{\sim}{a} = (a_1, a_2, \dots, a_n)$ coefficient vector (row vector)

$\underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ "variable" vector (column vector)

$\boxed{\text{Definition}}$ 1 $\underset{\sim}{a} \cdot \underset{\sim}{x} = a_1x_1 + a_2x_2 + \dots + a_nx_n$ i.e. $(a_1 \; a_2 \; \cdots \; a_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} a_ix_i$

so a single linear equation in n variables becomes $\underset{\sim}{a} \cdot \underset{\sim}{x} = b$.

suppose we require a system of 2 equations.

$$2x_1 + x_2 = 1$$
$$x_1 + x_2 = 3$$

2 unknowns

[if necessary we can separate them by a comma]

a_{ij} ⟶ equation i, variable j. $a_{1,11} \neq a_{11,1}$

we use double indices to represent coefficients in multi-equation systems.

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \dots + a_{mn}x_n = b_n$$

this is a system of <u>m linear equations</u> in <u>n variables</u>.

$\boxed{\text{Ex}}$ 1 Consider the following system of linear equations

$$\mathcal{S} = \begin{cases} x_1 + x_2 - x_3 = 1 \\ x_1 - x_2 + x_3 = 1 \\ x_1 \quad\; + 2x_3 = 3 \end{cases}$$

then

$a_{11}=1$ $a_{12}=1$ $a_{13}=-1$

$a_{21}=1$ $a_{22}=-1$ $a_{23}=1$

$a_{31}=1$ $a_{32}=0$ $a_{33}=2$

this notation was developed by Arthur Cayley, circa 1840.

Reconsidering the system S, where $S = \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots\cdots + a_{1n}x_n = b_1 \\ \vdots \qquad\qquad \ddots \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots\cdots + a_{mn}x_n = b_m \end{cases}$

then we can isolate the <u>coefficient matrix</u>, A and <u>variable vector</u>, $\underset{\sim}{x}$.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots\cdots & a_{mn} \end{pmatrix} \qquad \underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

where the first index is the <u>row index</u> → represents the equation no.

second index is the <u>column index</u> → corresponds to the vector it preceeds.

then $\boxed{A\underset{\sim}{x} = \underset{\sim}{b}}$, which is a matrix product.

recall that by Def. 1, for $\underset{\sim}{a} = (a_1, a_2, \cdots, a_n)$ and $\underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$,

then $\underset{\sim}{a} \cdot \underset{\sim}{x} = \sum_{i=1}^{n} a_i x_i = a_1 x_1 + \cdots + a_n x_n$.

now consider two matrices

$$A = \begin{pmatrix} a_{11} & \cdots\cdots & a_{1n} \\ a_{m1} & \cdots\cdots & a_{mn} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & \cdots\cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{N1} & \cdots\cdots & b_{Np} \end{pmatrix}$$

$m \times n$ matrix $\qquad N \times p$ matrix

$\boxed{\text{Definition}}$ 2 if and only if $n = N$ [i.e. AB is a sensible matrix product],

then if A is $m \times n$ and B is $n \times p$

AB will yield a $m \times p$ matrix, and $(AB)_{ij} = (i^{th} \text{ row of } A) \cdot (j^{th} \text{ column of } B)$ $\sum_{k=1}^{n} a_{ik} b_{kj}$

$\boxed{\text{Ex.}}$ 2 let $A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & -1 \\ 0 & 2 \\ -1 & 1 \end{pmatrix}$ $\Rightarrow$ AB is a 2×2 matrix.

$2 \times 3 \qquad 3 \times 2$

then $(AB)_{11} = (1^{st} \text{ row of } A) \cdot (1^{st} \text{ column of } B) = (1 \; 0 \; 1) \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 0$

$(AB)_{12} = (1^{st} \text{ row of } A) \cdot (2^{nd} \text{ column of } B) = (1 \; 0 \; 1) \cdot \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = 0$

$(AB)_{21} = (2^{nd} \text{ row of } A) \cdot (1^{st} \text{ column of } B) = (-1 \; 2 \; 3) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = -4$

$(AB)_{22} = (2^{nd} \text{ row of } A) \cdot (2^{nd} \text{ column of } B.) = (-1 \; 2 \; 3) \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = 8$

and so $AB = \begin{pmatrix} 0 & 0 \\ -4 & 8 \end{pmatrix}$

so, remember that AB is defined $\Leftrightarrow$ no. of columns of A = no. of rows of B.

$\boxed{\text{Ex.}}$ 2 (cont'd) since B is 3×2 and A is 2×3, thus BA is also defined and is a 3×3 matrix.

then $(BA)_{11} = (1 \; -1) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 2$

$(BA)_{12} = (1 \; -1) \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} = -2$ etc...... so $BA = \begin{pmatrix} 2 & -2 & -2 \\ -2 & 4 & 6 \\ -2 & 2 & 2 \end{pmatrix}$

$(BA)_{13} = (1 \; -1) \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} = -2$

note that AB is a 2×2 matrix, but BA is a 3×3 matrix.

ASPECTS OF MATRIX MULTIPLICATION.

① if A is $m \times n$ and B is $N \times p$ → (i) AB is defined when $n = N$, BA is defined when $p = m$.

(ii) AB is defined $\not\Rightarrow$ BA is defined.

1201-002.

(iii) if AB, BA both are defined, then may not be equal (or even the same size)

AB and BA are the same size

$\Leftrightarrow \quad m = n = p$

i.e. A and B are square matrices of the same size.

(iv)

even when A, B are both $n \times n$, which implies AB, BA are both defined,

usually $BA \neq AB$.

② matrix multiplication is <u>NON-COMMUTATIVE</u>.

$\boxed{\text{Ex.}}$ 3 $\quad A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

then $AB = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ and $BA = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

as can be seen, $AB \neq BA$ //

$\boxed{\text{Definition}}$ $\quad \forall\, m, n \in \mathbb{Z}^+ \quad \exists$ a <u>zero matrix</u> of size $m \times n$

For instance, if $m=2$ and $n=4$, $\quad \underline{0} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

thus 0 refers in linear algebra to any zero matrix which makes logical sense.

NILPOTENT MATRICES.

In normal arithmetic, $\quad a^2 = 0 \Rightarrow a = 0$

However in matrix algebra, $\quad X^2 = 0 \not\Rightarrow X = 0$.

For instance, $\quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $X \neq 0$

however $X^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and thus, $X^2 = 0$ but $X = 0$.

$\boxed{\text{Definition}}$ $\quad$ A <u>nilpotent matrix</u> X is one where $X^n = 0$ but $X \neq 0$.

IDENTITY MATRICES.

Let $n \in \mathbb{Z}^+$, $n > 1$,

$\exists$ an <u>identity matrix</u> I_n which has the property that...

if A is an $m \times n$ matrix and B is an $n \times p$ matrix,

$A I_n = A \qquad$ and $\qquad I_n B = B$.

hence, $\quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \ldots\ldots$

$\boxed{\text{Ex.}}$ 4 $\quad$ if $A = \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}$, (3×2 matrix)

then

$A I_2 = \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$= \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} = A$ // q.e.d.

if $B = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$, (2×3 matrix)

then

$I_2 B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$

$= \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = B$ // q.e.d.

$\boxed{\text{Definition}}$ $\quad$ we use the symbol δ_{ij} to denote the <u>kronecker delta</u>, where we fix n and $1 \leq i \leq n$, $1 \leq j \leq n$.

then $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ and this gives an $n \times n$ matrix I_n

For instance, when $n = 3$

$$\begin{matrix} \delta_{11} = 1 & \delta_{12} = 0 & \delta_{13} = 0 \\ \delta_{21} = 0 & \delta_{22} = 1 & \delta_{23} = 0 \\ \delta_{31} = 0 & \delta_{32} = 0 & \delta_{33} = 1 \end{matrix} \qquad \text{and} \qquad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

i.e. fixing $n \geq 1$, $\quad (I_n)_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$.

Proposition let $A = (A_{ij})$ $1 \le i \le m$, $1 \le j \le n$ be an $m \times n$ matrix.

then i) $AI_n = A$ and ii) $I_m A = A$

Proof: recall from the definition of AB that $(AB)_{ik} = \sum_{r=1}^{n} A_{ir} \cdot B_{rk}$,

so $(AI_n)_{ik} = \sum_{r=1}^{n} A_{ir} \cdot (I_n)_{rk} = \sum_{r=1}^{n} A_{ir} \delta_{rk}$

$= A_{ik}\delta_{kk} + \sum_{r \ne k}^{n} A_{ir}\delta_{rk}$

$= A_{ik}(1) + \sum_{r\ne k}^{n} A_{ir}(0) = A_{ik}$

since $(AI_n)_{ik} = A_{ik}$, it follows that $AI_n = A$ // q.e.d.

likewise, since $(BA)_{ik} = \sum_{r=1}^{n} B_{ir} \cdot A_{rk}$,

then $(I_m A)_{ik} = \sum_{r=1}^{n} (I_m)_{ir} \cdot A_{rk} = \sum_{r=1}^{n} \delta_{ir} A_{rk}$

$= \delta_{ii} A_{ik} + \sum_{r \ne k}^{n} \delta_{ir} A_{rk}$

$= (1) A_{ik} + \sum_{r\ne k}^{n} A_{rk}(0) = A_{ik}$

since $(I_m A)_{ik} = A_{ik}$, it follows that $I_m A = A$ // q.e.d.

6 october 2011
Prof FEA Johnson
Darwin LT.

For a system $S = \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$ we note that $A\underset{\sim}{x} = \underset{\sim}{b}$.

solving systems of linear equations — some can be written down immediately.

for instance, row reduced matrices.

Ex. take $A = \begin{pmatrix} 1 & 1 & 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ $\underset{\sim}{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$

the corresponding system $S = \begin{cases} x_1 + x_2 + 2x_4 + x_6 = 1 \\ x_3 - x_4 + 2x_6 = 1 \\ x_5 - 3x_6 = 1 \\ 0 = 0 \end{cases}$ — trivial equation.

steps for solution i) circle all the leading variables

ii) eliminate the circled variables by expressing them as sum of others.

thus,

$x_1 = 1 - x_2 - 2x_4 - x_6$
$x_2 = x_2$
$x_3 = 1 + x_4 - 2x_6$
$x_4 = x_4$
$x_5 = 1 + 3x_6$
$x_6 = x_6$.

and the solution matrix is $\begin{pmatrix} 1 & -x_2 & -2x_4 & -x_6 \\ & x_2 & & \\ 1 & & +x_4 & -2x_6 \\ & & x_4 & \\ 1 & & & +3x_6 \\ & & & x_6 \end{pmatrix}$ //

also called the CANONICAL SOLUTION.

note that this system has infinitely many solutions as x_2, x_4, x_6 are parameters that can take any real value.

recognising a row reduced matrix.

- $a_{11} = 1$ • $a_{i1} = 0$ for all $i > 1$ (i.e. first column only has a zero in the top-left corner).
- In row 2, first non-zero entry is 1, the rest of the column is 0.
- for each non-zero row, the first non-zero entry is 1.

Ex. $A = \begin{pmatrix} 1 & 2 & 0 & -1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 1 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ $\underset{\sim}{b} = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \end{pmatrix}$ $\Rightarrow$ $\begin{matrix} x_1 + 2x_2 - x_4 + 2x_7 = 1 \\ x_3 + x_4 + x_6 + 3x_7 = -1 \\ x_5 + x_6 + 4x_7 = 4 \end{matrix}$ $\Rightarrow$

$x_1 = 1 - 2x_2 + x_4 - 2x_7$
$x_2 = x_2$
$x_3 = -1 - x_4 - x_6 - 3x_7$
$x_4 = x_4$
$x_5 = 4 - x_6 - 4x_7$
$x_6 = x_6$
$x_7 = x_7$.

↑ row-reduced matrix

↓ canonical solution is expressed as a matrix

dim A = 4. x_1, x_3, x_5 are dependent (constrained) variables; x_2, x_4, x_6, x_7 are independent variables.

1201-004.

(REDUCED) ROW-ECHLON MATRICES.

Formation by elimination — 3 possible steps. The solution is unchanged if...

Notation.

→ Add equation j to equation i. ① Add λ(equation j) to equation i — $\mathcal{E}(i,j;\lambda)$

→ Multiply any equation by $\lambda \neq 0$ ② Multiply (equation i) by $\lambda \neq 0$. — $\mathcal{D}(i,\lambda)$ where $\lambda \neq 0$

→ Interchange order of equations. ③ swap (equation i) and (equation j) — $\mathcal{P}(i,j)$

$\boxed{\text{Ex.}}$ Take the system, $s = \begin{cases} 2x_1 + 3x_2 = 1 \\ x_1 + 2x_2 = -1 \end{cases}$

convert it to the form $A\underset{\sim}{x} = \underset{\sim}{b}$

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \quad \underset{\sim}{b} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

we then form the augmented matrix $(A|b) = \left(\begin{array}{cc|c} 2 & 3 & 1 \\ 1 & 2 & -1 \end{array}\right)$

Apply operations $\mathcal{E}(i,j;\lambda)$, $\mathcal{D}(i,\lambda)$ and $\mathcal{P}(i,j)$ until we get to row echlon form:

$$\left(\begin{array}{cc|c} 2 & 3 & 1 \\ 1 & 2 & -1 \end{array}\right) \xrightarrow{\mathcal{P}(1,2)} \left(\begin{array}{cc|c} 1 & 2 & -1 \\ 2 & 3 & 1 \end{array}\right) \xrightarrow{\mathcal{E}(2,1;-2)} \left(\begin{array}{cc|c} 1 & 2 & -1 \\ 0 & -1 & 3 \end{array}\right) \xrightarrow{\mathcal{D}(2,-1)} \left(\begin{array}{cc|c} 1 & 2 & -1 \\ 0 & 1 & -3 \end{array}\right) \xrightarrow{\mathcal{E}(1,2;-2)} \left(\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & -3 \end{array}\right) /\!/ .$$

$x_1 = 5$ // and $x_2 = -3$ //.

$\boxed{\text{Ex.}}$ solve the system s: $\begin{cases} x_1 + x_2 - x_3 - x_4 + x_5 = 1 \\ x_1 + x_2 + x_3 + 2x_4 - x_5 = -1 \\ x_1 + x_2 - 3x_3 - 4x_4 + 3x_5 = 3 \end{cases}$

then $(A|b) = \left(\begin{array}{ccccc|c} 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 2 & -1 & -1 \\ 1 & 1 & -3 & -4 & 3 & 3 \end{array}\right) \xrightarrow[\mathcal{E}(3,1;-1)]{\mathcal{E}(2,1;-1)} \left(\begin{array}{ccccc|c} 1 & 1 & -1 & -1 & 1 & 1 \\ 0 & 0 & 2 & 3 & -2 & -2 \\ 0 & 0 & -2 & -3 & 2 & 2 \end{array}\right) \xrightarrow{\mathcal{D}(2,\frac{1}{2})} \left(\begin{array}{ccccc|c} 1 & 1 & -1 & -1 & 1 & 1 \\ 0 & 0 & 1 & 3/2 & -1 & -1 \\ 0 & 0 & -2 & -3 & 2 & 2 \end{array}\right) \xrightarrow[\mathcal{E}(1,2;1)]{\mathcal{E}(3,2;2)} \left(\begin{array}{ccccc|c} 1 & 1 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$

thus,

$$\begin{aligned} x_1 &= \quad -x_2 - \tfrac{1}{2}x_4 \\ x_2 &= \quad x_2 \\ x_3 &= -1 \quad -\tfrac{3}{2}x_4 + x_5 \\ x_4 &= \quad x_4 \\ x_5 &= \quad x_5 \end{aligned}$$

and canonical solution is $\begin{pmatrix} -x_2 - \frac{1}{2}x_4 \\ x_2 \\ -1 - \frac{3}{2}x_4 + x_5 \\ x_4 \\ x_5 \end{pmatrix}$

matrix inversion.

To invert A of size 2×2, we use the augmented matrix $(A|I_2)$

we then reduce A to row-echlon form,

e.g. $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$

$$(A|I_2) = \left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right) \xrightarrow{\mathcal{E}(1,2;-2)} \left(\begin{array}{cc|cc} 0 & -1 & 1 & -2 \\ 1 & 2 & 0 & 1 \end{array}\right) \xrightarrow{\mathcal{P}(1,2)} \left(\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 0 & -1 & 1 & 2 \end{array}\right) \xrightarrow[\mathcal{D}(2,-1)]{\mathcal{E}(1,2;2)} \left(\begin{array}{cc|cc} 1 & 0 & 2 & -3 \\ 0 & 1 & -1 & 2 \end{array}\right) = (I_2|A^{-1})$$

$$A^{-1} = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} /\!/ .$$

11 October 2011
Prof FEA Johnson.
Damin H.

$\boxed{\text{Definition}}$ formal definition for "reduced row echlon"

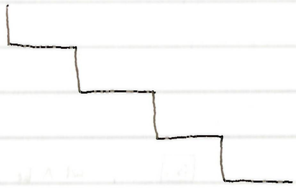
Let A be a matrix s.t. $A = (A_{ij})$ $1 \le i \le m$, $1 \le j \le n$ [i.e. matrix of size $m \times n$].

Let $(A_{i1}, A_{i2}, \dots, A_{in}) = A_{i*}$ = i^{th} row.

1^{st} condition: if $A_{i*} \neq 0$ and $A_{j*} = 0$, then $i < j$ (zero rows come last).

Let $A_{1*}, A_{2*}, \dots, A_{r*}$ be the non-zero rows,

let $A_{i,c(i)}$ be the first non-zero entry in A_{i*}.



2^{nd} condition: $A_{i,c(i)} = 1$

3^{rd} condition: $A_{k,c(i)} = 0$, $k \neq i$

4^{th} condition: $c(1) < c(2) < \dots < c(r)$

recall!

we solve systems of equations by reducing to row echlon form,

using three different operations

① $\mathcal{E}(i,j;\lambda)$
② $\mathcal{D}(i,\lambda)$ where $\lambda \neq 0$.
③ $\mathcal{P}(i,j)$

} elementary row operations.

we want to obtain the elementary row operations via matrix multiplication.

Ex.
$$\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} a+\lambda d & b+\lambda e & c+\lambda f \\ d & e & f \end{pmatrix}$$
$A \qquad\qquad A'$

We see that $A \xrightarrow{E(1,2;\lambda)} A'$.

$$\begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} a & b & c \\ \lambda a+d & \lambda b+e & \lambda c+f \end{pmatrix}$$
$A \qquad\qquad A''$

we see that $A \xrightarrow{E(2,1;\lambda)} A''$.

Hence, we see that the matrix for transformation $E(i,j;\lambda) = I_m + \lambda$ in $(i,j)^{th}$ position ? investigate......

Basic matrices -- only a single non-zero term.

For $m=2$,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$
$\epsilon(1,1) \quad \epsilon(1,2) \quad \epsilon(2,1) \quad \epsilon(2,2)$.

For $m=3$,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \cdots\cdots \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
$\epsilon(1,1) \quad \epsilon(1,2) \quad \epsilon(1,3) \quad \epsilon(2,1) \quad \epsilon(3,3)$.

Definition A basic matrix is one of the form $\epsilon(i,j)$, where $\epsilon(i,j)_{rs} = \delta_{ir}\delta_{js}$

$\Rightarrow \epsilon(i,j)_{rs} = \begin{cases} 1 & \text{if } r=i \text{ and } s=j \\ 0 & \text{otherwise} \end{cases}$.

Ex.
$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} a & b & c & d \\ e & f & g & h \\ k & l & m & n \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ k & l & m & n \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
A

this shows that $\epsilon(2,3)A$ causes all rows to be zero except row 2, and row 2 = old row 3.

Proposition Fix $m \geq 2$. Let A be $m \times n$. We then take $\epsilon(i,j)\cdot A$

Then (i^{th} row of $\epsilon(i,j))\cdot A$ = j^{th} row of A.

(k^{th} row of $\epsilon(i,j))\cdot A = 0$ if $k \neq i$.

i.e. take j^{th} row of A and put it in i^{th} row of product. Kill everything else.

Proof.
$$[\epsilon(i,j)A]_{rt} = \sum_{s=1}^{m} \epsilon(i,j)_{rs}\cdot A_{st} = \sum_{s=1}^{m} \delta_{ir}\delta_{js}\cdot A_{st}$$
$$= \delta_{ir}\ \delta_{jj}\ A_{jt} + \sum_{s\neq j} \delta_{ir}\delta_{js} A_{st}$$
$$= \delta_{ir} A_{jt} + 0 = \delta_{ir} A_{jt}$$

Hence $[\epsilon(i,j)\cdot A]_{rt} = \delta_{ir} A_{jt} = \begin{cases} A_{jt} & \text{if } r=i \\ 0 & \text{if } r\neq i \end{cases}$

$\Rightarrow$ i^{th} row of $\epsilon(i,j)\cdot A$ = j^{th} row of A, and

r^{th} row of $\epsilon(i,j)\cdot A = 0$ if $r \neq i$, q.e.d.

or alternatively $[\epsilon(i,j)\cdot A]_{r*} = \begin{cases} A_{j*} & r=i \\ 0 & r\neq i \end{cases}$.

Ex. Let A be a 3x3 matrix,

$$[\epsilon(2,3)\cdot A]_{1t} = \sum_{s=1}^{3} \epsilon(2,3)_{1s} A_{st}$$
$$= \epsilon(2,3)_{11}A_{1t} + \epsilon(2,3)_{12}A_{2t} + \epsilon(2,3)_{13}A_{3t} = \delta_{21}\delta_{31}A_{1t} + \delta_{21}\delta_{32}A_{2t} + \delta_{21}\delta_{33}A_{3t} = 0.$$

$[\epsilon(2,3)\cdot A]_{1t} = 0$ for all $t \in \mathbb{N}$ $\Rightarrow$ 1^{st} row of $\epsilon(2,3)\cdot A = 0$. (by the same logic, 3^{rd} row of $\epsilon(2,3)\cdot A = 0$).

$$[\epsilon(2,3)\cdot A]_{2t} = \sum_{s=1}^{3} \epsilon(2,3)_{2s}\cdot A_{st}$$
$$= \epsilon(2,3)_{21}A_{1t} + \epsilon(2,3)_{22}A_{2t} + \epsilon(2,3)_{23}\cdot A_{3t} = \delta_{22}\delta_{31}A_{1t} + \delta_{22}\delta_{32}A_{2t} + \delta_{22}\delta_{33}A_{3t} = A_{3t}$$

$\Rightarrow$ 2^{nd} row of $\epsilon(2,3)\cdot A$ = 3^{rd} row of A.

$\boxed{\text{Definition}}$ $E(i,j;\lambda) = I_m + \lambda \epsilon(i,j)$.

$\boxed{\text{Ex.}}$ where $m=4$; $E(2,3;5) = I_4 + 5\epsilon(2,3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Proposition: $E(i,j;\lambda)$ is the matrix obtained from A by performing operation $\mathcal{E}(i,j;\lambda)$.

i.e. i^{th} row of $E(i,j;\lambda)A$ = (i^{th} row of A) + λ(j^{th} row of A), while

k^{th} row of $E(i,j;\lambda)A$ = k^{th} row of A if $k \neq i$.

Proof — $E(i,j;\lambda)A = [I_m + \lambda\epsilon(i,j)]A$

$= I_m A + \lambda \epsilon(i,j) A = A + \lambda$ (the matrix where i^{th} row = j^{th} row of A, k^{th} row = 0 if $k \neq i$)

Hence, for the i^{th} row of the product,

i^{th} row of $I_mA + \lambda\epsilon(i,j)A$ = (i^{th} row of A) + λ(j^{th} row of A)

for $k \neq i$,

k^{th} row of $I_mA + \lambda\epsilon(i,j)A$ = (k^{th} row of A) + 0// q.e.d.

$\boxed{\text{Ex.}}$ let $m=4$, $A = \begin{pmatrix} a & b \\ c & d \\ e & f \\ g & h \end{pmatrix}$.

we multiply $E(2,3;5)$ before A,

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \\ e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a & b \\ c+5e & d+5f \\ e & f \\ g & h \end{pmatrix}//$$

$\boxed{\text{Proposition:}}$ $E(i,j;\lambda)\cdot E(i,j;\mu) = E(i,j;\lambda+\mu)$.

Proof — $(I+\lambda\epsilon(i,j))(I+\mu\epsilon(i,j)) = I^2 + \lambda\epsilon(i,j)I + I\mu\epsilon(i,j) + \epsilon(i,j)\epsilon(i,j)$

$= I + (\lambda+\mu)\epsilon(i,j) + \epsilon(i,j)^2$

$= I + (\lambda+\mu)\epsilon(i,j) + 0$

$= I + (\lambda+\mu)\epsilon(i,j) = E(i,j;\lambda+\mu)$// q.e.d.

[Proposition: (in homework) if $i \neq 0$, $\epsilon(i,j)\epsilon(i,j) = 0$. and more generally, $\epsilon(i,j)\epsilon(k,l) = \begin{cases} \epsilon(i,l) & \text{if } j=k \\ 0 & \text{if } j \neq k \end{cases}$]

note then, that for instance, $E(i,j;2)E(i,j;-2) = E(i,j;0) = I_n$

hence,

$\boxed{\text{Proposition}}$ $E(i,j;\lambda)$ is invertible, and $E(i,j;\lambda)^{-1} = E(i,j;-\lambda)$

Proof — $E(i,j;\lambda)\cdot E(i,j;-\lambda) = E(i,j;0) = I$// q.e.d.

$\therefore$ ①. operation of $\mathcal{E}(i,j;\lambda)$ on A = $E(i,j;\lambda)A$,

and $E(i,j;\lambda)$ is invertible.

we now examine $\mathcal{D}(i,\lambda)$:

$$\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} \xrightarrow{\mathcal{D}(2,5)} \begin{pmatrix} a & b \\ 5c & 5d \\ e & f \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix} = \begin{pmatrix} a & b \\ 5c & 5d \\ e & f \end{pmatrix}.$$

so it appears that $D(2,5)$ is transformed by.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (5-1)\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$\boxed{\text{Definition}}$ $D(i,\lambda) = I_m + (\lambda-1)\epsilon(i,i)$

$\boxed{\text{Proposition}}$ $D(i,\lambda)A$ is the matrix obtained from A by the operation $\mathcal{D}(i,\lambda)$.

recall that i^{th} row of $\epsilon(i,i)A$ = i^{th} row of A, $k \neq i$, k^{th} row of $\epsilon(i,i)A = 0$.

Proof — $D(i,\lambda)\cdot A = [I + (\lambda-1)\epsilon(i,i)]A = A + (\lambda-1)\epsilon(i,i)A$

by above, i^{th} row of $D(i,\lambda)A$ = i^{th} row of A + $(\lambda-1)$(i^{th} row of A) = λ(i^{th} row of A)

where $k \neq i$, k^{th} row of $D(i,\lambda)A$ = k^{th} row of A + $(\lambda-1)$(k^{th} row of A) = k^{th} row of A// q.e.d.

Proposition If $\lambda \neq 0$, $D(i,\lambda)$ is invertible and $D(i,\lambda)^{-1} = D(i,\frac{1}{\lambda})$.

Proposition $D(i,\lambda)\cdot D(i,\mu) = D(i,\lambda\mu)$.

Proof — $[I + (\lambda-1)\epsilon(i,i)][I + (\mu-1)\epsilon(i,i)] = I + (\lambda+\mu-2)\epsilon(i,i) + (\lambda-1)(\mu-1)\epsilon(i,i)^2$ — note: $\epsilon(i,i)^2 = \epsilon(i,i)$

$= I + (\lambda+\mu-2+\lambda\mu-\lambda-\mu+1)\epsilon(i,i)$

$= I + (\lambda\mu-1)\epsilon(i,i)$

$= D(i,\lambda\mu)$, q.e.d.

(Corollary) Proof — $D(i,\lambda)\cdot D(i,\frac{1}{\lambda}) = D(i,\lambda\cdot\frac{1}{\lambda}) = D(i,1) = I + (1-1)\epsilon(i,i) = I$, q.e.d.

$\therefore$ ②. operation of $D(i,\lambda) = D(i,\lambda)A$

and $D(i,\lambda)$ is reversible.

we now examine $P(i,j)$:

$P(i,j) \longrightarrow P(i,j) \leftarrow ?$ swapping i^{th} and j^{th} row.

guess: take $P(i,j)$ to be the matrix obtained by swapping i^{th} and j^{th} rows of I.

for instance, where $m=4$,

$$\begin{pmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1 \end{pmatrix} \xrightarrow{P(3,1)} \begin{pmatrix} 0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1 \end{pmatrix} \leftarrow \text{is this } P(3,1)?$$

we try,

$$\begin{pmatrix} 0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1 \end{pmatrix}\begin{pmatrix} a&b\\c&d\\e&f\\g&h \end{pmatrix} = \begin{pmatrix} e&f\\c&d\\a&b\\g&h \end{pmatrix}, \text{ which does work.}$$

Hence,

$\boxed{\text{Definition}}$ $P(i,j) = I_m - [\epsilon(i,i) + \epsilon(j,j)] + [\epsilon(i,j) + \epsilon(j,i)]$.

Proposition $P(i,j)A$ is the matrix obtained from A by swapping i^{th} and j^{th} rows.

Proof — $P(i,j)A = IA - \epsilon(i,i)A - \epsilon(j,j)A + \epsilon(i,j)A + \epsilon(j,i)A$

For the i^{th} row of A, it is $=$ i^{th} row of A $-$ i^{th} row of A $- 0 +$ j^{th} row of A $+0$

$=$ j^{th} row of A

For the j^{th} row of A, it is $=$ j^{th} row of A $- 0 -$ j^{th} row of A $+ 0 +$ i^{th} row of A

$=$ i^{th} row of A

For the k^{th} row of A, $k \neq i,j$, it is $=$ k^{th} row of A $-0-0+0+0 =$ k^{th} row of A, q.e.d.

Proposition $P(i,j)$ is invertible and $P(i,j)^{-1} = P(i,j)$

Proof $P(i,j)\cdot P(i,j) = [I - \epsilon(i,i) - \epsilon(j,j) + \epsilon(i,j) + \epsilon(j,i)]\cdot\begin{pmatrix} I\\ -\epsilon(i,i)\\ -\epsilon(j,j)\\ \epsilon(i,j)\\ \epsilon(j,i) \end{pmatrix}$

$= I - \epsilon(i,i) - \epsilon(j,j) + \epsilon(i,j) + \epsilon(j,i) - \epsilon(i,i) - \epsilon(j,j) - \epsilon(i,j) + \epsilon(i,i) + \epsilon(j,j) + \epsilon(i,j) - \epsilon(j,i)$
$+ \epsilon(i,i) + \epsilon(j,j) - \epsilon(i,j) + \epsilon(j,i) - \epsilon(j,i)$

$= I + 0 + 0 + 0 + 0 = I$, q.e.d.

FINDING A^{-1} with ELEMENTARY ROW OPERATIONS.

For example, let $A = \begin{pmatrix} 3&4\\2&3 \end{pmatrix}$.

Write $(A|I_m)$ and find Gaussian form on the left.

$$(A|I_m) = \left(\begin{array}{cc|cc} 3&4&1&0\\2&3&0&1 \end{array}\right) \xrightarrow{E(1,2;-1)} \left(\begin{array}{cc|cc} 1&1&1&-1\\2&3&0&1 \end{array}\right) \xrightarrow{E(2,1;-2)} \left(\begin{array}{cc|cc} 1&1&1&-1\\0&1&-2&3 \end{array}\right) \xrightarrow{E(1,2;-1)} \left(\begin{array}{cc|cc} 1&0&3&-4\\0&1&-2&3 \end{array}\right) = (I_m|A^{-1}).$$

verify that $AA^{-1} = I$ true, $\because \begin{pmatrix} 3&4\\2&3 \end{pmatrix}\begin{pmatrix} 3&-4\\-2&3 \end{pmatrix} = \begin{pmatrix} 1&0\\0&1 \end{pmatrix}$.

$(A|I_2) \longrightarrow E(1,2;-1)(A|I_m) \longrightarrow E(2,1;-2)\cdot E(1,2;-1)\cdot(A|I_m) \longrightarrow E(1,2;-1)\cdot E(2,1;-2)\cdot E(1,2;-1)(A|I_m) = (I_2|A^{-1})$.

How do we know that A^{-1} is produced on the right?

On the left, we see that $E(1,2;-1)\cdot E(2,1;-2)\cdot E(1,2;-1)A = I_2$, so by definition, since $A^{-1}A = I_m$, $A^{-1} = E(1,2;-1)\cdot E(2,1;-2)\cdot E(1,2;-1)$, q.e.d.

verifying numerically,

$$E(1,2;-1)\cdot E(2,1;-2)\cdot E(1,2;-1) = \begin{pmatrix}1&-1\\0&1\end{pmatrix}\begin{pmatrix}1&0\\-2&1\end{pmatrix}\begin{pmatrix}1&-1\\0&1\end{pmatrix}$$

$$= \begin{pmatrix}3&-1\\-2&1\end{pmatrix}\begin{pmatrix}1&-1\\0&1\end{pmatrix} = \begin{pmatrix}3&-4\\-2&3\end{pmatrix} = A^{-1} \;/\!/ \quad \text{(shown)}.$$

Hence, we have expressed A^{-1} as the product of elementary matrices.

Note also that since $(XY)^{-1} = Y^{-1}X^{-1}$ (because $Y^{-1}X^{-1}(XY) = Y^{-1}IY = I$ and $(XY)Y^{-1}X^{-1} = XIX^{-1} = I$).

then if $A^{-1} = X_1X_2X_3$, $A = X_3^{-1}X_2^{-1}X_1^{-1}$ //

$$= E(1,2;-1)^{-1}E(2,1;-2)^{-1}E(1,2;-1)^{-1}$$

$$= E(1,2;1)\,E(2,1;2)\,E(1,2;1)$$

$$= \begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}1&0\\2&1\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix} = \begin{pmatrix}3&1\\2&1\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix} = \begin{pmatrix}3&4\\2&3\end{pmatrix} \;/\!/ \quad \text{(verified)}.$$

18 October 2011
Prof FEA Johnson.
Darwin LT.

General approach to find A^{-1}:

Let A be an invertible $m\times m$ matrix.

Step 1: Form augmented matrix of size $m\times 2m$

$$(A|I_m)$$

Step 2: Proceed to reduce to row-echelon form

- Do this by left multiplication by suitable matrices $X_1, X_2, \dots, X_N$ in order.

$$(A|I_m) \xrightarrow{\text{oper. 1}} X_1(A|I_m) = (X_1A|X_1I_m) \xrightarrow{\text{oper. 2}} X_2(X_1A|X_1I_m) = (X_2X_1A|X_2X_1) \to \cdots$$

$$\xrightarrow{\text{oper. n}} X_N\cdots X_2X_1(A|I_m) = (X_N\cdots X_2X_1A|X_N\cdots X_2X_1).$$

- Finish when LHS becomes I_m.

since $X_N\cdots X_2X_1A = I$; by letting $B = X_N\cdots X_2X_1$,

$$BA = I \Rightarrow B = A^{-1}$$

thus $(X_N\cdots X_2X_1A|X_N\cdots X_2X_1) = (BA|B) = (I|B)$.

So this gives

i. an explicit representation for A^{-1} as a product of elementary invertible matrices, and also

ii. an explicit representation for A as a product

$$\because (A^{-1})^{-1} = A \Rightarrow$$

$$A = (X_N\cdots X_2X_1)^{-1}$$

$$= X_1^{-1}X_2^{-1}\cdots X_N^{-1}$$ ← note the reversal of order.

! order!

$\boxed{\text{Ex}}$ Find A^{-1}, where $A = \begin{pmatrix}2&0&1\\0&2&0\\1&0&1\end{pmatrix}$. Also express A^{-1} and A as a product of elementary matrices.

$$(A|I_m) = \left(\begin{array}{ccc|ccc}2&0&1&1&0&0\\0&2&0&0&1&0\\1&0&1&0&0&1\end{array}\right) \xrightarrow[E(3,1;-\frac{1}{2})]{D(1,\frac{1}{2}),\,D(2,\frac{1}{2})} \left(\begin{array}{ccc|ccc}1&0&\frac{1}{2}&\frac{1}{2}&0&0\\0&1&0&0&\frac{1}{2}&0\\0&0&\frac{1}{2}&-\frac{1}{2}&0&1\end{array}\right) \xrightarrow[D(3,2)]{E(1,3;-1)} \left(\begin{array}{ccc|ccc}1&0&0&1&0&-1\\0&1&0&0&\frac{1}{2}&0\\0&0&1&-1&0&2\end{array}\right) = (I_m|A^{-1}) \Rightarrow A^{-1} = \begin{pmatrix}1&0&-1\\0&\frac{1}{2}&0\\-1&0&2\end{pmatrix} /\!/$$

Hence, $A^{-1} = D(1,\frac{1}{2})\cdot D(2,\frac{1}{2})\cdot E(3,1;-\frac{1}{2})\cdot E(1,3;-1)\cdot D(3,2)$ // ✗

order should be reversed!!!

$$A = (A^{-1})^{-1} = [D(1,\tfrac{1}{2})\cdot D(2,\tfrac{1}{2})\,E(3,1;-\tfrac{1}{2})\,E(1,3;-1)\,D(3,2)]^{-1}$$

$$= D(3,2)^{-1}E(1,3;-1)^{-1}E(3,1;-\tfrac{1}{2})^{-1}D(2,\tfrac{1}{2})^{-1}D(1,\tfrac{1}{2})^{-1} = D(3,\tfrac{1}{2})E(1,3;1)E(3,1;\tfrac{1}{2})D(2,2)D(1,2) /\!/$$

PROPOSITIONAL LOGIC.

Logical keywords:

AND	OR	NOT	IMPLIES
$\wedge$	$\vee$	$\neg$	$\Rightarrow$

AND $\wedge$ (also called "conjunction")

For instance, let proposition p = "it is raining", and q = "it is cold".

Then we can construct a truth table as follows:

truth table (AND) :

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

OR $\vee$ ("inclusive OR")

truth table (OR) :

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

NOT $\neg$

truth table (NOT) :

p	$\neg p$
T	F
F	T

note that $\neg\neg p \equiv p$

notation : for logic, $A \equiv B$ means that A and B produce the same truth table.

also

p	q	$p \wedge q$	* $\neg(p \wedge q)$	$\neg p$	$\neg q$	* $\neg p \vee \neg q$
T	T	T	F	F	F	F
T	F	F	T	F	T	T
F	T	F	T	T	F	T
F	F	F	T	T	T	T

note from the table that $\neg(p \wedge q) \equiv \neg p \vee \neg q$

IMPLIES $\Rightarrow$

truth table (IMPLIES) :

p	q	$p \Rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

for instance, either "it is not cold" or "I will go out".

p	q	$\neg p$	$\neg p \vee q$
T	T	F	T
T	F	F	F
F	T	T	T
F	F	T	T

letting p = "it is cold", q = "I will go out"

Definition : "$p \Rightarrow q$" means exactly $\boxed{\text{"}\neg p \vee q\text{"}}$.

Why is this reasonable?

we start with $\neg p \vee p$, which is always true :

$\neg p$	p	$\neg p \vee p$
T	F	T
F	T	T

the idea of "$p \Rightarrow q$" is that if p, then q.

$1 \to \neg p$, $1 \to p \to q$.

since $\neg p \vee p$ is always true, we can say "either $\neg p$ or p" but "if p, then q",

and that boils down to "either $\neg p$ happens or q happens"

so "$p \Rightarrow q$" $\equiv$ "$\neg p \vee q$".

standard identities involving $\wedge$, $\vee$, $\neg$, $\Rightarrow$

1) $\neg\neg p \equiv p$

2) $(p \vee q) \vee r \equiv p \vee (q \vee r)$ 2') $(p \wedge q) \wedge r \equiv p \wedge (q \wedge r)$ [associativity]

3) $p \vee q \equiv q \vee p$ 3') $p \wedge q \equiv q \wedge p$ [commutativity]

4) $(p \vee q) \wedge r \equiv (p \wedge r) \vee (q \wedge r)$ 4') $(p \wedge q) \vee r \equiv (p \vee r) \wedge (q \vee r)$ [distributivity].

5) $\neg(p \wedge q) \equiv (\neg p) \vee (\neg q)$ 5') $\neg(p \vee q) \equiv (\neg p) \wedge (\neg q)$.

6) $\neg q \Rightarrow \neg p \equiv p \Rightarrow q$

Definition: $\neg p \Rightarrow \neg q$ is known as the <u>contrapositive</u> of $p \Rightarrow q$ and is equivalent to it

$q \Rightarrow p$ is known as the <u>converse</u> of $p \Rightarrow q$ and is different from it.

Note that $q \Rightarrow p \not\equiv p \Rightarrow q$

p	q	$p \Rightarrow q$	$q \Rightarrow p$
T	T	T	T
T	F	F	T
F	T	F	F
F	F	T	T.

$\boxed{\text{Ex.}}$ Is $q \Rightarrow (p \Rightarrow q)$ true?

p	q	$p \Rightarrow q$	$q \Rightarrow (p \Rightarrow q)$
T	T	T	T
T	F	F	T
F	T	T	T
F	F	T	T

Hence $q \Rightarrow (p \Rightarrow q)$ is always true.

A statement which is always T is called a <u>tautology</u>,

a statement which is always F is called a <u>contradiction</u>.

For instance, $p \wedge \neg p$

p	$\neg p$	$p \wedge \neg p$
T	F	F
F	T	F

Most compound statements are neither tautologies nor contradictions.

For instance,

p	q	r	$p \vee q$	$(p \vee q) \Rightarrow r$	$p \Rightarrow r$	$[(p \vee q) \Rightarrow r] \Rightarrow [p \Rightarrow r]$
T	T	T	T	T	T	T
T	T	F	T	F	F	T
T	F	T	T	T	T	T
T	F	F	T	F	F	T
F	T	T	T	T	T	T
F	T	F	T	F	T	T
F	F	T	F	T	T	T
F	F	F	F	T	T	T

$\Rightarrow$ tautology.

20 October 2011
Prof FEA Johnson
Parvin LT

$\boxed{\text{Ex}}$ contrast $[(p \vee q) \Rightarrow r] \Rightarrow [p \Rightarrow r]$ to $[(p \wedge q) \Rightarrow r] \Rightarrow [p \Rightarrow r]$.

p	q	r	$p \wedge q$	$(p \wedge q) \Rightarrow r$	$p \Rightarrow r$	$[(p \wedge q) \Rightarrow r] \Rightarrow [p \Rightarrow r]$
T	T	T	T	T	T	T
T	T	F	T	F	F	T
T	F	T	F	T	T	T
T	F	F	F	T	F	F
F	T	T	F	T	T	T
F	T	F	F	T	T	T
F	F	T	F	T	T	T
F	F	F	F	T	T	T

T $\frac{7}{8}$ of the time,
$\Rightarrow$ but not always
$\to$ not a tautology.

So what exactly are tautologies + contradictions?

- a tautology is a LOGICALLY CORRECT ARGUMENT.

Consider the following arguments:

(a) If it is cold <u>and</u> wet then the elephant will not dance. However it is not wet and the elephant is dancing.

$\Rightarrow$ it is not cold. $\boxed{\text{FALSE}}$

(b) If it is cold <u>or</u> wet then the elephant will not dance. However it is not wet and the elephant is dancing.

$\Rightarrow$ it is not cold.

Let p = it is cold, q = it is wet, r = the elephant is not dancing.

(a) $([(p \wedge q) \Rightarrow r] \wedge [\neg q \Rightarrow \neg r]) \Rightarrow \neg p$

(b) $([(p \vee q) \Rightarrow r] \wedge [\neg q \Rightarrow \neg r]) \Rightarrow \neg p$.

p	q	r	$p \wedge q$ OR	$p \vee q$	S_A: $(p \wedge q) \Rightarrow r$ OR	S_B: $(p \vee q) \Rightarrow r$	T: $\neg q \wedge \neg r$	$S_A \wedge T$ OR	$S_B \wedge T$	$(S_A \wedge T) \Rightarrow \neg p$ or $(S_B \wedge T) \Rightarrow \neg p$
T	T	T	T	T	T	T	F	F	F	T
T	T	F	T	T	F	F	F	F	F	T
T	F	T	F	T	T	T	F	F	F	T
T	F	F	F	T	T	F	T	T	F	T
F	T	T	F	T	T	T	F	F	F	T
F	T	F	F	T	T	F	F	F	F	T
F	F	T	F	F	T	T	F	F	F	T
F	F	F	F	F	T	T	T	T	T	T

thus, (a) is contingent but (b) is a tautology.

History of Truth Tables — a lineage:

C.S Pierce $\longrightarrow$ L. Wittgenstein $\longrightarrow$ A. Turing.

Note that not all 4 symbols are necessary!

$p \wedge q \equiv \neg\neg p \wedge \neg\neg q \equiv \neg(\neg p \vee \neg q)$ while $p \vee q \equiv \neg(\neg p \wedge \neg q)$.

likewise, $p \Rightarrow q \equiv \neg p \vee q$.

$p \vee q \equiv \neg p \Rightarrow q$.

} only two signs are actually necessary: $\vee, \neg$ or $\wedge, \neg$.

In fact, just one symbol suffices; using Sheffer's stroke function $a|b$, where $|$ implies that a,b are not <u>both</u> true.

p	q	p|q	p|(p|q)
T	T	F	T
T	F	T	F
F	T	T	T
F	F	T	T

. thus, $p|(p|q) \equiv p \Rightarrow q$

p	p|p
T	F
F	T.

LOGIC OF VARIABLE EXPRESSIONS (Predicate Logic).

Think of expressions which vary just between 0 or 1.

For example, $P(x)$ is "$x=0$" where x is either 0 or 1.

Then $P(0)$ is true, and $P(1)$ is false.

The possible range of variation of x is called the domain $\mathcal{D}$.

starting with $\mathcal{D} = \{0,1\}$.

From propositional calculus, we take $\wedge, \vee, \Rightarrow, \neg$; as well as $\forall$ (universal quantifier) and $\exists$ (existence symbol/ Existential quantifier).

$(\forall x)\, P(x)$ = <u>for all</u> $x \in \mathcal{D}$, $P(x)$ is T.

$\exists x\, P(x)$ = <u>for at least one</u> $x \in \mathcal{D}$, $P(x)$ is T.

Questions to consider:

- How do $\forall, \exists$ interact with $\wedge, \vee, \neg, \Rightarrow$?
- can we interpret $\forall, \exists$ in terms of Propositional calculus?

For instance, if $x = 0$ or 1, $\forall x\, Q(x)$ means both $Q(0)$ is true and $Q(1)$ is true.

so for $\mathcal{D} = \{0,1\}$, $(\forall x) Q(x) \equiv Q(0) \wedge Q(1)$; likewise $(\exists x) Q(x) \equiv Q(0) \vee Q(1)$.

25 October 2011
Prof FEA Johnson.
Darwin LT.

Imagine if $\mathcal{D} = \{0,1,2\}$, then $\forall x\, P(x) \equiv P(0) \wedge P(1) \wedge P(2)$

then if $\mathcal{D} = \mathbb{N} \cup \{0\} = \{0,1,2,\ldots\}$, then $\forall x\, P(x) \equiv P(0) \wedge P(1) \wedge P(2) \ldots \wedge P(n) \wedge P(n+1) \wedge \ldots$

↪ this is the reason why we adopt a symbol "$\forall$".

if $\mathcal{D} = \{0,1,2\}$, then $\exists x\, P(x) \equiv P(0) \vee P(1) \vee P(2)$.

then if $\mathcal{D} = \mathbb{N} \cup \{0\} = \{0,1,2,\ldots\}$, then $\exists x\, P(x) \equiv P(0) \vee P(1) \vee P(2) \ldots \vee P(n) \vee P(n+1) \vee \ldots$

For a domain of infinite length, $\forall$ behaves like a repeated $\wedge$ (logical product), $\exists$ behaves like a repeated $\vee$ (logical sum).

Negation of Quantifiers.

Take $\mathcal{D} = \{0,1\}$; then $\forall x\, P(x) \equiv P(0) \wedge P(1)$.

$\neg(\forall x)\, P(x) \equiv \neg[P(0) \wedge P(1)] \equiv \neg P(0) \vee \neg P(1) \equiv (\exists x)\, \neg P(x)$.

$\Rightarrow$ we take it axiomatically that $\boxed{\neg(\forall x)\, P(x) \equiv (\exists x)\, \neg P(x)}$

likewise, since $(\exists x) P(x) \equiv P(0) \vee P(1)$

$\neg(\exists x)\, P(x) \equiv \neg[P(0) \vee P(1)] \equiv \neg P(0) \wedge \neg P(1) \equiv (\forall x)\, \neg P(x)$.

$\Rightarrow$ we take it axiomatically that $\boxed{\neg(\exists x)\, P(x) \equiv (\forall x)\, \neg P(x)}$

1201-012

Interchange of order of quantifiers

ⓐ $(\forall x)(\forall y)\,P(x,y) \equiv (\forall y)(\forall x)\,P(x,y)$

ⓑ $(\exists x)(\exists y)\,P(x,y) \equiv (\exists y)(\exists x)\,P(x,y)$

these are equivalent statements – use one to prove the other.

Suppose we assume ⓐ. Then NTP ⓐ ⇒ ⓑ:

$(\exists x)(\exists y)P(x,y) \equiv (\exists x)(\exists y)\neg\neg P(x,y) \equiv (\exists x)\neg(\forall y)\neg P(x,y) \equiv \neg(\forall x)(\forall y)\neg P(x,y) \overset{\text{using ⓐ}}{\equiv} \neg(\forall y)(\forall x)\neg P(x,y)$

$\equiv (\exists y)(\exists x)\neg\neg P(x,y) \equiv (\exists y)(\exists x)P(x,y)$ // proven.

Hence, the order is unimportant when quantifiers are the same.

However, $(\exists x)(\forall y)$ is not the same as $(\forall y)(\exists x)$

for instance, take $D=\{0,1\}$ and $P(x,y)$ be "$x=y$".

$(\forall x)(\exists y)P(x,y) \equiv (\forall x)\{P(x,0) \vee P(x,1)\} \equiv [P(0,0) \vee P(0,1)] \wedge [P(1,0) \vee P(1,1)]$

$\equiv [T \vee F] \wedge [F \vee T] \equiv T \wedge T \equiv T.$

Hence $(\forall x)(\exists y)P(x,y)$ is true; but flipping the order,

$(\exists y)(\forall x)\,P(x,y) \equiv (\exists y)\{P(0,y) \wedge P(1,y)\} \equiv [P(0,0) \wedge P(1,0)] \vee [P(0,1) \wedge P(1,1)]$

$\equiv [T \wedge F] \vee [F \wedge T] \equiv F \vee F \equiv F.$

thus, $(\exists y)(\forall x)P(x,y)$ is false.

as such, it follows that $(\forall x)(\exists y)\,P(x,y) \not\equiv (\exists y)(\forall x)P(x,y)$;

so in general, $(\forall x)(\exists y)P \not\Rightarrow (\exists y)(\forall x)P$.

SET THEORY

An <u>ordered</u> set is one in which order is important — there is a first term, second term etc.

Hence, $(0,1) \neq (1,0)$.

In set theory, order is not important as <u>unordered</u> sets are considered.

$\{0,1,2\} \neq \{1,2,0\}$; in contrast to $(0,1,2) \neq (1,2,0)$

For small, finite sets, elements can be listed out individually. But for large sets? Infinite sets?

The symbol '$\in$' means belongs to, and thus '$x \in A$' means 'x is a member of set A'.

Large sets need to be described by their "defining property".

for instance, $X = \{2n : n \in \mathbb{N}\}$, informally $\{2,4,6,\dots\}$

$Y = \{n \in \mathbb{N} : n \geq 3\}$, informally $\{3,4,5,\dots\}$

→ here, "$n \geq 3$" is the <u>defining predicate</u> of set Y.

$Z = \{n \in \mathbb{N},\ (n \geq 3) \wedge (n \leq 10)\}$, informally $\{3,4,5,6,7,8,9,10\}$.

Sets are defined by their elements — what comes within the curly brackets, ignoring order.

general form of a set A is

$$A = \{\underset{\text{typical element}}{x} : \underset{\text{defining predicate}}{P_A(x)}\}$$

so if $E(x)$ means "x is an elephant" and $P(x)$ means "x is pink",

then for $A = \{x : E(x) \wedge P(x)\}$, we claim that:

A is the set of pink elephants.

$\boxed{\text{Definition}}$ Let A be a set with defining predicate $P_A(x)$, and B be a set with defining predicate $P_B(x)$;

we say that B is a <u>subset</u> of A [written $B \subset A$] when $P_B(x) \Rightarrow P_A(x)$.

or, informally, if $x \in B$ then $x \in A$.

$\boxed{\text{Ex}}$ Let $A = \{0,1,2,3,4,5,6\}$ and $B = \{1,3,5\}$ *Notation: DO NOT CONFUSE "$\subset$" and "$\in$"

then $B \subset A$, because every element in B is simultaneously an element in A

$\boxed{\text{Ex}}$ Given a set $A = \{0,1,2\}$ and $B = \{1,2\}$, then $B \subset A$.

Now take $C = \{0,1,2,\{1,2\}\}$

Be careful – here C has only 4 elements, not 5!

thus, we can say $2 \in C$, $\{1,2\} \subset C$ and $\{1,2\} \in C$. These are all true.

subset with 1,2 is in C the single element $\{1,2\}$ is in C

but while $\{0,1\} \subset C$ is true; $\{0,1\} \in C$ is false!

subset with 0,1 is in C no such element exists in C.

$\boxed{\text{Ex}}$ Take $D = \{0,1,\{0,1\},\{1,2\}\}$. Then we have the following statements:

$\{0,1\} \in D$, $\{0,1\} \subset D$, $\{1,2\} \in D$, $\{1,2\} \not\subset D$, $\{\{0,1\},\{1,2\}\} \subset D$.

the elements 1,2 are not in D.

$1 \in D$, $\underline{1 \not\subset D}$, $\{1\} \notin D$, $\{1\} \subset D$

notation error: "1" is not a set, and so obviously cannot be a subset!

Predicates of sets.

Let A, B be sets, then

the <u>union</u> of A and B, $\underline{A \cup B}$, consists of those elements which are in A or in B.

$$\Rightarrow P_{A\cup B}(x) \equiv P_A(x) \vee P_B(x)$$

$$A \cup B = \{x : P_A(x) \vee P_B(x)\}.$$

$\boxed{\text{Ex}}$ $A = \{0,1,2\}$ and $B = \{2,3,4\}$, then $A \cup B = \{0,1,2,3,4\}$.

the <u>intersection</u> of A and B, $\underline{A \cap B}$, consists of those elements which are in A *and* in B.

$$\Rightarrow P_{A\cap B}(x) \equiv P_A(x) \wedge P_B(x)$$

$$A \cap B = \{x : P_A(x) \wedge P_B(x)\}.$$

$\boxed{\text{Ex}}$ $A = \{0,1,2\}$ and $B = \{2,3,4\}$, then $A \cap B = \{2\}$

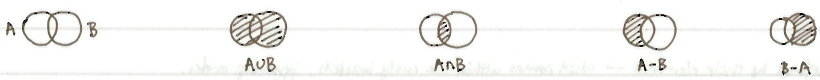
the <u>difference</u> of A and B, $\underline{A - B}$ or $\underline{A \setminus B}$, consists of those elements which are in A *but not* in B.

$$\Rightarrow P_{A-B}(x) \equiv P_A(x) \wedge \neg P_B(x)$$

$$A - B = \{x : P_A(x) \wedge \neg P_B(x)\}$$

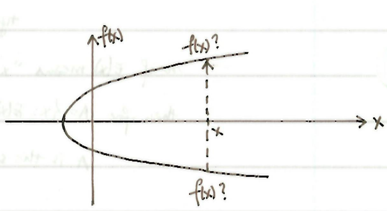
*note that unlike $\cup$ and $\cap$, $\setminus$ is not commutative.

these can be expressed by Venn diagrams, but these diagrams are useful only for up to 3 sets when represented on a plane.



MAPPINGS / FUNCTIONS (the two words are interchangeable)

In early Mathematics, Newton considered curves (circa 1665), such as the ones on the right.



In response to the inability to determine a unique $f(x)$ for x,

N.H. Abel introduced the notion of functions.

1201-014

The idea of a function is that...

Two sets A, B and a "rule" f, which associates to each $x \in A$ a single well defined element $f(x) \in B$.

then $f: A \to B$.
domain codomain

27 October 2011
Prof FEA Johnson
Darwin LT.

A function/mapping must consist of a

(i) set A, the domain

(ii) set B, the codomain

(iii) a rule which associates each $a \in A$ a single well-defined element $f(a) \in B$.

for it to be considered a "function"

Remember: a function is not the same as a formula.

Examples of functions → $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x+1$ $\quad$ $f: \mathbb{R}-\{-1\} \to \mathbb{R}$, $f(x) = \frac{1}{x+1}$ (note that codomain need not exclude 0).

$f: \mathbb{R}-\{\pm 1\} \to \mathbb{R}$, $f(x) = \frac{x^4-1}{x^2-1}$ $\quad$ $f: \mathbb{R}_- \cup \{0\} \to \mathbb{R}_+ \cup \{0\}$, $f(x) = -\sqrt{x}$.

consider the following –

let $g: \mathbb{R} \to \mathbb{R}$, $g(x) = x^2+1$

$h: \mathbb{R} \to \mathbb{R}$, $h(x) = \begin{cases} \frac{x^4-1}{x^2-1} & \text{if } x \neq \pm 1 \\ x^2+1 & \text{if } x = \pm 1 \end{cases}$

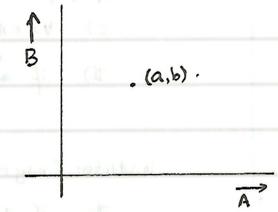
Here, g and h have different formulae, but nevertheless represent the same function.

Products of sets.

The idea of $A \times B$ is that a point in each space can be represented by coordinates (a,b) such that $a \in A$ and $b \in B$.

given ordered pairs (a,b), (a',b').

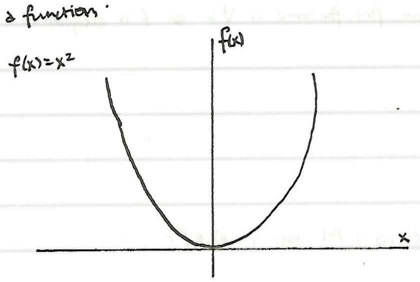
the rule for equality states that $(a,b) = (a',b') \iff a=a'$ and $b=b'$.



$\boxed{\text{Definition}}$ The product of sets A and B,

$A \times B = \{(a,b) : `a \in A' \wedge `b \in B'\}$.

Graph of a function.



let $f: \mathbb{R} \to \mathbb{R}_+$.

the the graph of f is given by $\{(x, f(x)) : x \in \mathbb{R}\}$.

in this case this represents a subset of $\mathbb{R} \times \mathbb{R}_+$.

$\boxed{\text{Definition}}$ Let A, B be sets.

By a mapping/function $f: A \to B$, we mean a subset $f \subset A \times B$ satisfying the following conditions:

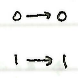
i) $\forall a \in A$, $\exists b \in B$ s.t. $(a,b) \in f$. $\quad$ hence $f(a)$ is defined for all $a \in A$.

[when $(a,b) \in f$ we write $f(a) = b$]

ii) $(a,b) \in f$ and $(a,b') \in f \Rightarrow b = b'$ $\quad$ hence $f(a)$ is a single element.

$\boxed{\text{Ex}}$ Describe all functions $\alpha: \{0,1\} \to \{1,0\}$.

i) $\alpha(0)=0$, $\alpha(1)=1$

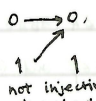


injective
surjective.

ii) $\alpha(0)=1$, $\alpha(1)=0$

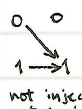
injective
surjective

iii) $\alpha(0)=0$, $\alpha(1)=0$



not injective
not surjective

iv) $\alpha(0)=1$, $\alpha(1)=1$.



not injective
not surjective

Definition A mapping $f: A \to B$ is said to be injective when $f(a) = f(a') \Rightarrow a = a'$.

Likewise, its contrapositive: $a \neq a' \Rightarrow f(a) \neq f(a')$.

Definition Let $f: A \to B$ be a mapping.

We say that f is surjective when $\forall b \in B$, $\exists a \in A$ s.t. $f(a) = b$.

Ex Take $f: \mathbb{R} \to \mathbb{R}$. Is f injective and/or surjective? $f(x) = x^2$.

NOT injective : $f(1) = f(-1)$

NOT surjective : $-1 \in \mathbb{R}$ cannot be attained.

Ex Take $g: \mathbb{N} \to \mathbb{N}$. Is g injective and/or surjective? $g(x) = 2x$.

injective : $g(x_1) = g(x_2) \Rightarrow x_1 = x_2$.

NOT surjective : $1 \in \mathbb{N}$ cannot be attained

Ex Take $h: \mathbb{R} \to \mathbb{R}$. Is h injective and/or surjective? $h(x) = 2x$.

injective : $h(x_1) = h(x_2) \Rightarrow x_1 = x_2$.

surjective: all $\mathbb{R}$ can be attained as $h(\frac{y}{2}) = y$.

1 November 2011
Prof FEA Johnson.
Damin LT.

Formally, if A, B are sets, then

a mapping $f: A \mapsto B$ is a subset $f \subset A \times B$ such that the following two properties are defined.

I) $\forall a \in A$, $\exists b \in B$; $(a, b) \in f$. (then we write $b = f(a)$).

II) if $a \in A$, $b, b' \in B$ are such that $(a, b) \in f$ and $(a, b') \in f$, then $b = b'$ (i.e. $f(a)$ is uniquely defined).

In addition, f may or may not have the following properties.

III) if $a, a' \in A$, $b \in B$ are st. $(a, b) \in f$ and $(a', b) \in f$, then $a = a'$

i.e. $f(a) = f(a') \Rightarrow f$ is injective.

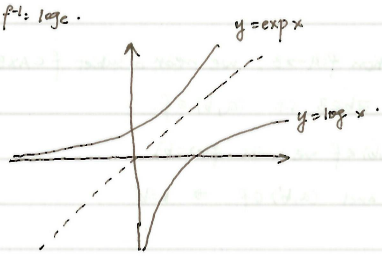
IV) $\forall b \in B$, $\exists a \in A$; $(a, b) \in f$. (i.e. b has form $f(a)$ for some a $\forall b \Rightarrow f$ is surjective).

In general a subset $f \subset A \times B$ need not have any of these properties.

A general subset $f \subset A \times B$ is called a relation.

Definition if $f \subset A \times B$ then $f^{-1} \subset B \times A$ is the set $\{(b, a) \in f^{-1} \Leftrightarrow (a, b) \in f\}$.

Ex If $f = \exp$ then $f^{-1}: \log_e$.



$\exp: \mathbb{R} \to \mathbb{R}^+ = \{z \in \mathbb{R}, z > 0\}$.

$\log: \mathbb{R}^+ \to \mathbb{R}$

Question: let $f: A \to B$ be a mapping. When is f^{-1} also a mapping?

$f \subset A \times B$.	$f^{-1} \subset B \times A$.
(I) $\forall a \in A$ $\exists b \in B$: $(a, b) \in f$.	(I') $\forall b \in B$ $\exists a \in A$: $(b, a) \in f^{-1}$.
(II) $(a, b) \in f$ and $(a, b') \in f \Rightarrow b = b'$.	(II') $(b, a) \in f^{-1}$ and $(b, a') \in f^{-1} \Rightarrow a = a'$. $(a, b) \in f$ and $(a', b) \in f \Rightarrow a = a'$.

Observe that.
(III) for f (injective) $\equiv$ (II) for f^{-1}
(IV) for f (surjective) $\equiv$ (I) for f^{-1}.

1201-016.

$\boxed{\text{Corollary}}$ If $f: A \mapsto B$ is a mapping and is both injective and surjective, then $f^{-1}: B \mapsto A$ is a mapping $\therefore$

$\underline{f \subset A \times B}$, $\underline{f^{-1} \subset B \times A}$.

(II) $(a,b) \in f$ and $(a',b) \in f \Rightarrow a = a'$ (III)' $(b,a) \in f^{-1}$ and $(b',a) \in f^{-1} \Rightarrow b = b'$.
$(a,b) \in f$ and $(a,b') \in f \Rightarrow b = b'$.

(III) $\forall b \in B$ $\exists a \in A$ $(a,b) \in f$. (IV)' $\forall a \in A$ $\exists b \in B$: $(b,a) \in f^{-1}$ = $(a,b) \in f$.

observe too that:
(III) for $f^{-1} \equiv$ (II) for f
(IV) for $f^{-1} \equiv$ (I) for f.

$\boxed{\text{Definition}}$ A mapping is <u>bijective</u> if it is both injective and surjective.

complete $\boxed{\text{Corollary}}$: Let $f \subset A \times B$ be a subset. Then f is a bijective mapping $\Leftrightarrow$ f^{-1} is a bijective mapping.

COMPOSITION OF MAPPINGS.

Let A, B, C be sets, and let $f: A \mapsto B$ and $g: B \mapsto C$ be mappings.

Define $g \circ f: A \mapsto C$ by $(g \circ f)(a) = g[f(a)]$ *first f then g: order of composition makes a difference!

$\boxed{\text{Ex}}$ $f: \mathbb{R} \to \mathbb{R}$ $g: \mathbb{R} \to \mathbb{R}$.
$f(x) = x^2$ $g(x) = \sin(x) + 1$
$g \circ f(x) = \sin x^2 + 1$ $f \circ g(x) = (\sin(x) + 1)^2$.
note that $g \circ f \neq f \circ g$; show by taking a counterexample e.g. $x = \pi$.
then $f \circ g(\pi) = 1$, but $g \circ f(\pi) = \sin \pi^2 + 1 \neq 1 = f \circ g(\pi)$.

$\boxed{\text{Ex}}$ $f(x) = \exp(x)$ $f: \mathbb{R} \mapsto \mathbb{R}^+$; $g(x) = \log(x)$ $g: \mathbb{R}^+ \mapsto \mathbb{R}$.
$g \circ f(x) = x$ and $f \circ g(x) = x$, but $g \circ f \neq f \circ g$! because
$g \circ f: \mathbb{R} \mapsto \mathbb{R}$ but $f \circ g: \mathbb{R}^+ \mapsto \mathbb{R}^+$. $\Rightarrow$ $g \circ f = Id_{\mathbb{R}}: \mathbb{R} \mapsto \mathbb{R}$, $f \circ g = Id_{\mathbb{R}^+}: \mathbb{R}^+ \mapsto \mathbb{R}^+$

$\boxed{\text{Definition}}$ If A is a set, the <u>identity mapping</u> on A, Id_A is given such that.
$Id_A: A \mapsto A$, $Id_A(a) = a$ $\forall a \in A$.

$\boxed{\text{Definition}}$ A mapping $f: A \mapsto B$ is said to be <u>invertible</u> when $\exists$ a mapping $g: B \mapsto A$ s.t.
$g \circ f = Id_A$ and $f \circ g = Id_B$ (note the different domains!).

$\boxed{\text{Ex}}$ $\exp: \mathbb{R} \to \mathbb{R}^+$ is invertible i.e. $\log \circ \exp = Id_{\mathbb{R}}$.
likewise,
$\log: \mathbb{R}^+ \to \mathbb{R}$ is invertible i.e. $\exp \circ \log = Id_{\mathbb{R}^+}$.

The composition of mappings is associative:

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$$

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Proof — calculate, for some $a \in A$,

$$\left.\begin{aligned} h \circ (g \circ f)(a) &= h[(g \circ f)(a)] = h(g(f(a))) \\ (h \circ g) \circ f(a) &= h \circ g(f(a)) = h(g(f(a))) \end{aligned}\right\} \text{LHS} = \text{RHS}.$$

$\boxed{\text{Proposition}}$ If $f: A \mapsto B$ is an invertible mapping, then its inverse is unique.

Proof — suppose $(g \circ f) = Id_A$ and $(h \circ f) = Id_A$; $(f \circ g) = Id_B$ and $(f \circ h) = Id_B$, we claim $h = g$.
$h \circ (f \circ g) = h \circ Id_B = h$ $\Rightarrow$ $h \circ Id_B(b) = h(b)$; $(h \circ f) \circ g = Id_A \circ g = g$ but $h \circ (f \circ g) = (h \circ f) \circ g$ $\Rightarrow$ $h = g$ // q.e.d.

Putting everything together...

$\boxed{\text{Theorem}}$ If $f: A \mapsto B$ is a bijective mapping, then f is invertible with inverse f^{-1}.

Proof — f is bijective $\Rightarrow$ f^{-1} is a mapping.

clearly $f \circ f^{-1} = \text{Id}$, $f^{-1} \circ f = \text{Id}$ // q.e.d.

Thus far, we have shown that f is bijective $\Rightarrow$ f is invertible.

conversely,

$\boxed{\text{Proposition}}$ Let $f: A \mapsto B$ be an invertible mapping, then f is bijective.

Proof — we know $\exists$ mapping $g: B \to A$ s.t. $g \circ f = \text{Id}_A$, $f \circ g = \text{Id}_B$.

f is injective: suppose $f(a) = f(a')$.

then $g \circ f(a) = g \circ f(a')$ $\Rightarrow$ $\text{Id}_A(a) = \text{Id}_A(a')$ $\Rightarrow$ $a = a'$, so $a = a'$.

f is surjective: let $b \in B$, we need to find $a \in A$ s.t. $f(a) = b$.

take $a = g(b)$, then $f(a) = f(g(b)) = b$ $\Rightarrow$ $f \circ g = \text{Id}$ // q.e.d.

Hence, the complete result is that...

$\boxed{\text{Theorem}}$ If $f: A \mapsto B$ is a mapping, then f is invertible $\Leftrightarrow$ f is bijective.

$\boxed{\text{Ex}}$ $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ is not injective, $\because$ $f(1) = f(-1)$ but $1 \neq -1$.

in this case, we can make f injective by restricting domain; we now get.

$f:$ $\mathbb{R}^+ \setminus \{0\} \to \mathbb{R}$. is injective, but it is not surjective. $\because$ $-1 \in \mathbb{R}$, codomain, but $\nexists$ $x \in \mathbb{R}^+$ s.t. $x^2 = -1$.

in this case, we can make f surjective by restricting codomain; we now get.

$f: \mathbb{R}^+ \cup \{0\} \to \mathbb{R}^+ \cup \{0\}$. now f is bijective, which makes it invertible, and:

$f^{-1}: \mathbb{R}^+ \cup \{0\} \longrightarrow \mathbb{R}^+ \cup \{0\}$, $f(y) = \sqrt{y}$.

PERMUTATIONS.

A mapping $f: \{1, 2, \dots, n\} \longrightarrow \{1, 2, \dots, n\}$ is said to be a permutation on n $\Leftrightarrow$ f is bijective.

equivalently, $\Leftrightarrow$ f is invertible.

For instance, where $n=2$, all mappings $\{1,2\} \to \{1,2\}$ include:

$1 \to 1$, $2 \to 2$	$1 \to 1$, $2 \to 1$	$1 \to 2$, $2 \to 2$	$1 \to 2$, $2 \to 1$
Id (bijective!)	constant to 1	constant to 2.	τ where $\tau(1) = 2$, $\tau(2) = 1$ (bijective!)

hence there are 2 permutations of $\{1,2\}$, $\text{Id} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$, $\tau = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

where $n=3$, all mappings $\{1,2,3\} \longrightarrow \{1,2,3\}$ include: all 27 possibilities. however, for bijective ones, there are only $6 = 3!$

$\text{Id} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ $x = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ $x^2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ $xy = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ $x^2y = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ $y = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$.

hence there are 6 permutations of $\{1,2,3\}$.

let $x = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, then $x^2 = x \circ x = \left.\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}\begin{matrix} \\ \text{by } x \\ \text{by } x \end{matrix}\right. = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$.

is $yx = xy$? NO. $yx = y \circ x = \left.\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 1 & 3 & 2 \end{pmatrix}\begin{matrix} \\ \text{by } x \\ \text{by } y \end{matrix}\right. = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = x^2y$.

$x^3 = \text{Id}$ $y^2 = \text{Id}$ $yx = x^2y$

in general, there are $n!$ permutations of $\{1, 2, \dots, n\} \longrightarrow \{1, 2, \dots, n\}$.

Convention: if $f: \{1, 2, \dots, n\} \longrightarrow \{1, 2, \dots, n\}$ is a mapping,

then we write $f = \begin{pmatrix} 1 & 2 & 3 & \cdots\cdots & n \\ f(1) & f(2) & f(3) & \cdots\cdots & f(n) \end{pmatrix}$.

CYCLIC PERMUTATIONS.

Take $\{a_1, \dots, a_m\} \subset \{1, 2, \dots, n\}$.

$\boxed{\text{Definition}}$ $(a_1, a_2, \dots, a_m)$ is defined to be the permutation which takes...

$a_1 \to a_2,\ a_2 \to a_3,\ \dots,\ a_{m-1} \to a_m,\ a_m \to a_1$; and leaves everything else fixed.

$\boxed{\text{EX}}$ let $n=10$, $m=5$.

$(a_1, a_2, a_3, a_4, a_5) = (2, 7, 5, 3, 9)$.

then $(2,7,5,3,9) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 1 & 7 & 9 & 4 & 3 & 6 & 5 & 8 & 2 & 10 \end{pmatrix}$

shifted by 1 —— do not change.

we say $(a_1, a_2, \dots, a_m)$ is a cycle of length m, then $(a_1, a_2, \dots, a_m)^m = \mathrm{id}$.

convention: $f = \begin{pmatrix} 1 & 2 & \cdots\cdots & n \\ f(1) & f(2) & \cdots\cdots & f(n) \end{pmatrix}$.

Proposition — If f, g are both permutations on n letters then so also is $g \circ f$.

Proof — $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$; and since f, g are permutations $\Longleftrightarrow$ f, g are invertible, then $g \circ f$ is invertible, q.e.d.

Composition of permutations,

$g \circ f = \begin{pmatrix} 1 & 2 & 3 & \cdots\cdots & n \\ f(1) & f(2) & f(3) & \cdots\cdots & f(n) \\ g(1) & g(2) & g(3) & \cdots\cdots & g(n) \end{pmatrix}$ ← cancel middle row $= \begin{pmatrix} 1 & 2 & 3 & \cdots\cdots & n \\ gf(1) & gf(2) & gf(3) & \cdots\cdots & gf(n) \end{pmatrix}$.

Formal statement of a cyclic permutation.

$(a_1, a_2, \dots, a_m)(r) = \begin{cases} a_{i+1} & \text{if } r = a_i,\ i < m \\ a_1 & \text{if } r = a_m \\ r & \text{if } r \notin \{a_1, a_2, \dots, a_m\} \end{cases}$.

usually, but not always, permutations do not commute.

Exceptional circumstances: Let $\{a_1, a_2, \dots, a_m\} \subset \{1, 2, \dots, n\}$ and $\{b_1, b_2, \dots, b_k\} \subset \{1, 2, \dots, n\}$

① suppose $\{a_1, a_2, \dots, a_m\} \cap \{b_1, b_2, \dots, b_k\} = \emptyset$ then $(a_1, a_2, \dots, a_m)\cdot(b_1, b_2, \dots, b_k) = (b_1, b_2, \dots, b_k)(a_1, a_2, \dots, a_m)$.

DISJOINT CYCLES COMMUTE.

$\boxed{\text{EX}}$ $n=8$, then $(2,5,7)(1,6,4) = (1,6,4)(2,5,7)$.

$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 2 & 3 & 1 & 5 & 4 & 7 & 8 \\ 6 & 5 & 3 & 1 & 7 & 4 & 2 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 2 & 3 & 1 & 5 & 4 & 7 & 8 \\ 6 & 5 & 3 & 1 & 7 & 4 & 2 & 8 \end{pmatrix}$ // (verified).

Rule 0: Disjoint cycles commute

Rule 1: Any permutation is a product of disjoint cycles. (and from rule 0, they commute)

$\boxed{\text{EX}}$ Let $\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 5 & 10 & 7 & 8 & 11 & 2 & 12 & 4 & 3 & 6 & 1 & 9 \end{pmatrix} = (1,5,11)(2,10,6)(3,7,12,9)(4,8)$. commutative!

$\boxed{\text{Definition}}$ By a <u>transposition</u>, we mean a cycle of length 2.

i.e. it switches the two elements i and j, and leaves everything else fixed, hence a transposition has form (i, j).

Rule 2: Any cycle is a product of transpositions. // a cycle of length m is a product of $m-1$ transpositions.

$\boxed{\text{EX}}$ Let $m=4$, $m \le n$, then imagine a cycle (a_1, a_2, a_3, a_4).

(a_1,a_2,a_3,a_4):

$a_1 \cdots a_2 \cdots a_3 \cdots a_4 \cdots$

$a_2 \cdots a_1 \cdots a_3 \cdots a_4 \cdots \; \} (a_1, a_2)$.

$a_2 \cdots a_3 \cdots a_1 \cdots a_4 \cdots \; \} (a_1, a_3)$

$a_2 \cdots a_3 \cdots a_4 \cdots a_1 \cdots \; \} (a_1, a_4)$

so $(a_1, a_2, a_3, a_4) = (a_1, a_4)(a_1, a_3)(a_1, a_2)$.

cycle of length 4. product of 3 transpositions.

generalisation —— $(a_1, a_2, \dots, a_m) = (a_1, a_m)(a_1, a_{m-1}) \cdots\cdots (a_1, a_3)(a_1, a_2)$.

$\boxed{\text{Definition}}$ An <u>adjacent transposition</u> is one of the form $(i, i+1)$.

Rule 3: Any transposition is a product of an odd number of adjacent transformations. (gap of k ~ $2k-1$ adjacent transformations.

Proof — WLOG let $i<j$, then we define $\text{gap}(i,j) = j-1$.

an adjacent transposition is a transposition of gap 1.

suppose we consider $(i, i+k)$, $k>1$. ; then $i < i+1 < i+k$.

$i \quad i+1 \ \cdots \ i+k$

$i+1 \quad i \ \cdots \ i+k \; \} \text{gap}(i, i+1) = 1$

$i+k \quad i \quad i+1 \; \} \text{gap}(i+1, i+k) = k-1$

$i+k \quad i \; \} \text{gap}(i, i+1) = 1$.

we claim that a transposition of gap k is a product of $2k-1$ adjacent transpositions. For $k=1$, nothing to prove.

assume $k = m-1$ is true, then $k = m$ is true as well because $1 + (2k-3) + 1 = 2k-1$.

if σ is a permutation,

$$\text{sign}(\sigma) = \begin{cases} +1 & \text{if } \sigma \text{ is a product of an EVEN no. of adjacent transpositions} \\ -1 & \text{if } \sigma \text{ is a product of an ODD no. of adjacent transpositions} \end{cases}$$

then,

sign (cycle of even length) $= -1$

sign (cycle of odd length) $= +1$

$\boxed{\text{Ex}}$ $\sigma = (1,5,11)(2,10,6)(3,7,12,9)(4,8)$.

3	3	4	2	lengths.
+1	+1	-1	-1	signs

$\Rightarrow \text{sign}(\sigma) = +1$.

$\boxed{\text{Definition}}$ $\text{order}(\sigma) = \min\{r \geq 1 ; \sigma^r = \text{Id}\}$.

15 November 2011.
Prof FEA Johnson.
Darwin LT.

we describe $\mathbb{Q}$ as the set of numbers where $\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\}$.

rule of equality : $\frac{p}{q} = \frac{p'}{q'} \iff pq' = p'q$.

$\mathbb{Q}$ is a system in which one can add, subtract, multiply and divide $(\neq 0)$

Any system which has these properties is called a field

$\boxed{\text{Definition}}$ A field $\mathbb{F}$ consists of the following data.

$$\mathbb{F} = (\mathbb{F}, +, 0, \cdot, 1)$$

where $\mathbb{F}$ is a set.

$0, 1 \in \mathbb{F}$; $0 \neq 1$

$+$ is a mapping $\quad +: \mathbb{F} \times \mathbb{F} \longmapsto \mathbb{F}$

$(x,y) \longmapsto x+y$ [we write this rather than $+(x,y)_{\text{polish}}$.] } addition.

and $\cdot$ is also a mapping $\quad \cdot: \mathbb{F} \times \mathbb{F} \longmapsto \mathbb{F}$

$(x,y) \longmapsto x \cdot y$ } multiplication.

such that the following properties hold

1) $x+(y+z) = (x+y)+z \quad \forall x,y,z \in \mathbb{F}$ (associativity) $\qquad$ 1)' $x\cdot(y\cdot z) = (x\cdot y)\cdot z \quad \forall x,y,z \in \mathbb{F}$.

2) $x+y = y+x \quad \forall x,y \in \mathbb{F}$ (commutativity). $\qquad$ 2)' $x\cdot y = y\cdot x \quad \forall x,y \in \mathbb{F}$.

3) $x+0 = x \quad \forall x \in \mathbb{F}$ (additive identity). $\qquad$ 3)' $x\cdot 1 = x \quad \forall x \in \mathbb{F}$ (multiplicative identity)

4) $\forall x \in \mathbb{F} \ \exists (-x) \in \mathbb{F} ; \ x+(-x) = 0$ (additive inverse). $\qquad$ 4)' $\forall x \in \mathbb{F}\setminus\{0\}, \ \exists x^{-1} \in \mathbb{F}$ s.t. $xx^{-1} = 1$. (multiplicative inverse).

5) $(x+y)\cdot z = x\cdot z + y \cdot z$; $z\cdot(x+y) = zx \cdot zy$ (distributivity).

so $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are fields. However, $\mathbb{Z}$ is not a field as 4)' does not hold.

Let $\mathbb{F}_2$ be a field with two elements ; e.g. $\mathbb{F}_2 = \{0,1\}$.

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

we let 0,1 represent even and odd numbers respectively.

so $\mathbb{F}_2$ is obtained from $\mathbb{Z}$ by forcing '$2=0$' i.e '$1+1=0$'.

by extension, we examine $\mathbb{F}_3 = \{0,1,2\}$; introduce rule '$3=0$'.

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

i.e. $2^{-1} = 2$ in $\mathbb{F}_3$.

now we try $\mathbb{F}_5 = \{0,1,2,3,4\}$

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3.

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

However, note that this does not apply for $\{0,1,2,3\}$ where '$4=0$'

·	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

this is not a field, because 2^{-1} does not exist.
so this is not what we call $\mathbb{F}_4$!

Let $\mathbb{F}$ be a field.

$\boxed{\text{Definition}}$ $\mathbb{F}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \underset{\sim}{x} \ ; \ x_i \in \mathbb{F} \right\}$.

where $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \iff \forall i \ \ x_i = y_i$

$\mathbb{F}^n$ has the following properties:

$+ : \mathbb{F}^n \times \mathbb{F}^n \longmapsto \mathbb{F}^n$.

$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{pmatrix}$ i.e. adding coordinates.

$\cdot : \mathbb{F} \times \mathbb{F}^n \longmapsto \mathbb{F}^n$ note that we multiply here only by a scalar $\lambda \in \mathbb{F}$; not a vector $\mathbb{F}^n$.

$\lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}$

there exists a zero vector $\underline{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ which has the following properties.

additive.

1) $\underline{x} + (\underline{y} + \underline{z}) = (\underline{x} + \underline{y}) + \underline{z}$ $(\underline{x}, \underline{y}, \underline{z} \in \mathbb{F}^n)$

2) $\underline{x} + \underline{y} = \underline{y} + \underline{x}$ $(\underline{x}, \underline{y} \in \mathbb{F}^n)$

3) $\underline{x} + \underline{0} = \underline{x}$ $(\underline{x} \in \mathbb{F}^n)$

4) $\forall \underline{x} \in \mathbb{F}^n \ \exists (-\underline{x}) \in \mathbb{F}^n$ s.t. $\underline{x} + (-\underline{x}) = \underline{0}$

multiplicative.

5) $\lambda \cdot (\mu \cdot \underline{z}) = (\lambda \cdot \mu) \cdot \underline{z}$ $(\forall \lambda, \mu \in \mathbb{F}; \ \forall \underline{z} \in \mathbb{F}^n)$

6) $1 \cdot \underline{z} = \underline{z}$ $(\forall \underline{z} \in \mathbb{F}^n)$

distributive laws.

7) $(\lambda + \mu) \cdot \underline{z} = \lambda \cdot \underline{z} + \mu \cdot \underline{z}$ $(\forall \lambda, \mu \in \mathbb{F}; \ \forall \underline{z} \in \mathbb{F}^n)$.

8) $\lambda \cdot (\underline{w} + \underline{z}) = \lambda \cdot \underline{w} + \lambda \cdot \underline{z}$ $(\forall \lambda \in \mathbb{F}; \ \forall \underline{w}, \underline{z} \in \mathbb{F}^n)$

Definition Let $\mathbb{F}$ be a field. Suppose that

i) V is a set

ii) $\underline{0} \in V$

iii) $+ : V \times V \longmapsto V$ is a mapping

iv) $\cdot : \mathbb{F} \times V \longmapsto V$ is a mapping

then we say that $(V, +, 0, \cdot)$ is a <u>vector space</u> over $\mathbb{F}$ when the analogous properties 1) – 8) hold.

or, restating it...

Definition Let $\mathbb{F}$ be a field; then $V = (V, +, \underline{0}, \cdot)$ is a <u>vector space over $\mathbb{F}$</u>.

when i) V is a set and $\underline{0} \in V$

ii) $+ : V \times V \mapsto V$ is a mapping such that:

a) $\underline{x} + (\underline{y} + \underline{z}) = (\underline{x} + \underline{y}) + \underline{z}$

b) $\underline{x} + \underline{y} = \underline{y} + \underline{x}$

c) $\underline{x} + \underline{0} = \underline{x}$

d) $\exists \underline{x} \in V, \ \exists (-\underline{x}) \in V : \underline{x} + (-\underline{x}) = \underline{0}$.

iii) $\cdot : \mathbb{F} \times V \mapsto V$ is a mapping such that:

e) $\lambda \cdot (\mu \cdot \underline{z}) = (\lambda \cdot \mu) \cdot \underline{z}$

f) $1 \cdot \underline{z} = \underline{z}$

g) $(\lambda + \mu) \cdot \underline{z} = \lambda \underline{z} + \mu \underline{z}$

h) $\lambda (\underline{w} + \underline{z}) = \lambda \underline{w} + \lambda \underline{z}$

Ex $\mathbb{F}^n$ is a vector space over $\mathbb{F}$; in the special case where $n = 1$, we regard $\mathbb{F} = \mathbb{F}^1$, so $\mathbb{F}$ is also a vector space.

Ex (more typical example); take $\mathbb{F} = \mathbb{Q}$; and take $V = \left\{ \begin{pmatrix} a \\ -a \end{pmatrix} : a \in \mathbb{Q} \right\}$; so $V \subset \mathbb{Q}^2$

addition in V: $\begin{pmatrix} a \\ -a \end{pmatrix} + \begin{pmatrix} b \\ -b \end{pmatrix} = \begin{pmatrix} a+b \\ -(a+b) \end{pmatrix} \in V$

scalar multiplication: $\lambda \begin{pmatrix} a \\ -a \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda(-a) \end{pmatrix} \in V$

so we have addition $+ : V \times V \mapsto V$; scalar multiplication $\cdot : \mathbb{Q} \times V \longmapsto V$; and a zero vector $\underline{0} = \begin{pmatrix} 0 \\ -0 \end{pmatrix} \in V$.

all the other axioms hold because they already hold in $\mathbb{Q}^2$; and $V \subset \mathbb{Q}^2$.

however; $V \neq \mathbb{Q}^2$ — use counter-example $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{Q}^2$; $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \notin V$.

Ex Take $\mathbb{F}$ to be any field; let $W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{F}^3 : x_1 + x_2 + x_3 = 0 \right\}$.

W has an addition $\underline{x}, \underline{y} \in W$; $\underline{x} + \underline{y} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ x_3+y_3 \end{pmatrix}$ $(x_1+y_1)+(x_2+y_2)+(x_3+y_3) = 0 \Rightarrow \underline{x} + \underline{y} \in W$.

W has scalar multiplication $\underline{x} \in W$; $\lambda \in \mathbb{F}$; then $\lambda \underline{x} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \end{pmatrix}$ $\lambda x_1 + \lambda x_2 + \lambda x_3 = \lambda(0) = 0 \Rightarrow \lambda \underline{x} \in W$.

remaining axioms already hold $\because$ they hold for $\mathbb{F}_3$, and $W \subset \mathbb{F}_3$.

however; $W \neq \mathbb{F}^3$ — use counter-example $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \notin W$.

DIMENSIONALITY OF A VECTOR SPACE.

we examine the system $\mathbb{F}^n$.

we define $\underline{e}_1 = \begin{pmatrix}1\\ \vdots\\ 0\end{pmatrix}$, $\underline{e}_2 = \begin{pmatrix}0\\1\\ \vdots\\ 0\end{pmatrix}$, $\ldots$, $\underline{e}_n = \begin{pmatrix}0\\0\\ \vdots\\ 1\end{pmatrix}$

for instance, when $n=4$,

$$\underline{e}_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad \underline{e}_2 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad \underline{e}_3 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad \underline{e}_4 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

$\underline{e}_1, \underline{e}_2, \underline{e}_3, \underline{e}_4$ in $\mathbb{F}^4$ have the following property:

imagine $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{F}$; then suppose $\lambda_1\underline{e}_1 + \lambda_2\underline{e}_2 + \lambda_3\underline{e}_3 + \lambda_4\underline{e}_4 = \underline{0}$, then the only possibility is that $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$.

to see this, note that $\lambda_1\underline{e}_1 + \lambda_2\underline{e}_2 + \lambda_3\underline{e}_3 + \lambda_4\underline{e}_4 = \begin{pmatrix}\lambda_1\\0\\0\\0\end{pmatrix} + \begin{pmatrix}0\\\lambda_2\\0\\0\end{pmatrix} + \begin{pmatrix}0\\0\\\lambda_3\\0\end{pmatrix} + \begin{pmatrix}0\\0\\0\\\lambda_4\end{pmatrix} = \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\\\lambda_4\end{pmatrix}$

so if $\lambda_1\underline{e}_1 + \cdots + \lambda_4\underline{e}_4 = \underline{0}$; then $\begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\\\lambda_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}$.

LINEAR INDEPENDENCE

$\boxed{\text{Definition}}$ If $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n \in V$, where V is a vector space over $\mathbb{F}$; $\sum_{i=1}^{n}\lambda_i\underline{v}_i = \underline{0}$

we state that $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ is <u>linearly independent</u> when $\lambda_1\underline{v}_1 + \cdots + \lambda_n\underline{v}_n = \underline{0} \Rightarrow \lambda_1 = \cdots = \lambda_n = 0$

i.e. the only way for a linear combination of $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ to sum to $\underline{0}$ is if all coefficients are 0.

17 october 2011
Prof FEA Johnson.
Darwin LT.

An expression $\lambda_1\underline{v}_1 + \lambda_2\underline{v}_2 + \cdots + \lambda_n\underline{v}_n$ $(\lambda_i \in \mathbb{F})$ is called a <u>linear combination</u>.

A set of vectors $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ is linearly independent if the only linear combination which is $\underline{0}$, is the obvious one.

$\boxed{\text{Ex}}$ we consider $\{\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n\} \in \mathbb{F}^n$; then $\{\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n\}$ is linearly independent.

Proof – take $\sum_{i=1}^{n}\lambda_i\underline{e}_i = \begin{pmatrix}\lambda_1\\\lambda_2\\\vdots\\\lambda_n\end{pmatrix}$

so if $\sum\lambda_i\underline{e}_i = \underline{0}$, then $\lambda_i = 0$ for each $i \leq n$, q.e.d.

$\boxed{\text{Ex}}$ we take the vectors $\underline{\varepsilon}_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}$, $\underline{\varepsilon}_2 = \begin{pmatrix}1\\1\\0\end{pmatrix}$ and $\underline{\varepsilon}_3 = \begin{pmatrix}1\\0\\0\end{pmatrix}$; then $\{\underline{\varepsilon}_1, \underline{\varepsilon}_2, \underline{\varepsilon}_3\}$ is linearly independent.

Proof – $\lambda_1\underline{\varepsilon}_1 = \begin{pmatrix}\lambda_1\\\lambda_1\\\lambda_1\end{pmatrix}$; $\lambda_2\underline{\varepsilon}_2 = \begin{pmatrix}\lambda_2\\\lambda_2\\0\end{pmatrix}$; $\lambda_3\underline{\varepsilon}_3 = \begin{pmatrix}\lambda_3\\0\\0\end{pmatrix}$.

so $\lambda_1\underline{\varepsilon}_1 + \lambda_2\underline{\varepsilon}_2 + \lambda_3\underline{\varepsilon}_3 = \begin{pmatrix}\lambda_1+\lambda_2+\lambda_3\\\lambda_1+\lambda_2\\\lambda_1\end{pmatrix}$

and hence, $\sum_{i=1}^{3}\lambda_i\underline{\varepsilon}_i = \underline{0} \Rightarrow \lambda_1+\lambda_2+\lambda_3 = 0$; $\lambda_1+\lambda_2 = 0$; $\lambda_1 = 0 \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0$, q.e.d.

A set which is not linearly independent is said to be <u>linearly dependent</u>.

$\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ is LD $\Longleftrightarrow$ $\exists$ a linear combination $\sum_{i=1}^{n}\lambda_i\underline{v}_i = \underline{0}$ where $\lambda_i \neq 0$ for at least one value of $i \leq n$.

$\boxed{\text{Ex}}$ we take the vectors $\underline{w}_1 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$, $\underline{w}_2 = \begin{pmatrix}1\\-1\\-1\\-1\end{pmatrix}$, $\underline{w}_3 = \begin{pmatrix}1\\2\\3\\0\end{pmatrix}$, $\underline{w}_4 = \begin{pmatrix}1\\4\\3\\2\end{pmatrix}$; then $\{\underline{w}_1, \underline{w}_2, \underline{w}_3, \underline{w}_4\}$ is LD.

Proof (counter-example) – $\underline{w}_4 = \underline{w}_1 - \underline{w}_2 + \underline{w}_3 \Rightarrow \underline{w}_1 - \underline{w}_2 + \underline{w}_3 - \underline{w}_4 = 0$.

$\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = 1$, $\lambda_4 = -1$; so in this case at least one coefficient is non-zero.

Standard mistake! "every set is LI because $0\underline{v}_1 + 0\underline{v}_2 + \cdots + 0\underline{v}_n = \underline{0}$ and all coefficients are 0." WRONG.

LI means the <u>only</u> way to get 0 is where $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

SPANNING.

If $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$ are standard vectors in $\mathbb{F}^n$, and $\underline{x} \in \mathbb{F}^n$ is an arbitrary vector,

$\underline{x} = \begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$ and then $\underline{x} = x_1\underline{e}_1 + x_2\underline{e}_2 + \cdots + x_n\underline{e}_n$.

for instance, $\underline{x} = \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = x_1\underline{e}_1 + x_2\underline{e}_2 + x_3\underline{e}_3$.

$\boxed{\text{Definition}}$ If $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n \in V$; we say that $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ <u>span</u> V when any vector $\underline{x} \in V$ can be written in the form

$\underline{x} = \lambda_1\underline{v}_1 + \cdots + \lambda_n\underline{v}_n$ for some $\lambda_i \in \mathbb{F}$ $= \sum_{i=1}^{n}\lambda_i\underline{v}_i$.

so we have seen that the standard vectors $\underline{e}_1, \underline{e}_2, \ldots, \underline{e}_n$ span $\mathbb{F}^n$.

$\boxed{\text{Definition}}$ Let $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ be vectors in V. we say that $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ is a <u>basis</u> for V when

i) $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ is linearly independent and ii) $\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\}$ spans V.

Ex The standard vectors $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \dots, \underset{\sim}{e}_n\}$ form a basis for $\mathbb{F}^n$.

Theorem BASIS THEOREM.

i) Any non-zero vector space V has a basis.

ii) Any two bases for V have the same number of elements. *Proof: see 1201-031.

Definition The dimension $\dim(V)$ of V is the number of elements in a basis for V.

From the above, we observe that $\dim(\mathbb{F}^n) = n$. Also, we call $\{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \dots, \underset{\sim}{e}_n\}$ a <u>standard basis</u> for V over $\mathbb{F}^n$.

Ex Let $W = \{ \underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{F}^3 : x_1 + x_2 + x_3 = 0 \}$; then $\dim W = 2$.

we need to find a <u>basis</u> for W with exactly 2 elements.

take $\varphi_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ and $\varphi_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$; $\varphi_1, \varphi_2 \in W$.

to show that φ_1, φ_2 is a basis, it must fulfill the two conditions of the basis theorem.

- $\{\varphi_1, \varphi_2\}$ is LI.

  $\lambda_1 \varphi_1 + \lambda_2 \varphi_2 = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ -(\lambda_1 + \lambda_2) \end{pmatrix}$, so if $\lambda_1 \varphi_1 + \lambda_2 \varphi_2 = \underset{\sim}{0}$; then $\lambda_1 = \lambda_2 = 0$// q.e.d.

- $\{\varphi_1, \varphi_2\}$ spans W.

  take $\underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in W$; then $x_3 = -(x_1 + x_2)$.

  then $x_1 \varphi_1 + x_2 \varphi_2 = \begin{pmatrix} x_1 \\ 0 \\ -x_1 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ -x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ -(x_1 + x_2) \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underset{\sim}{x}$.

hence, $\dim W = 2$//.

Ex Let A be a $m \times n$ matrix $(a_{ij} \in \mathbb{F})$

Define $K_A = \{ \underset{\sim}{x} \in \mathbb{F}^n ; A\underset{\sim}{x} = 0 \}$ i.e. the solution set to homogeneous system of linear equations.

Proposition — K_A is a vector space $\subset \mathbb{F}^n$.

Proof — if $\underset{\sim}{x}, \underset{\sim}{y} \in K_A$; then

$A\underset{\sim}{x} = 0,\ A\underset{\sim}{y} = 0 \Rightarrow A(\underset{\sim}{x} + \underset{\sim}{y}) = A\underset{\sim}{x} + A\underset{\sim}{y} = \underset{\sim}{0}$; so $\underset{\sim}{x} + \underset{\sim}{y} \in K_A$. (addition).

$x \in K_A,\ \lambda \in \mathbb{F} \Rightarrow A(\lambda \underset{\sim}{x}) = \lambda(A\underset{\sim}{x}) = \lambda \underset{\sim}{0} = \underset{\sim}{0}$; so $\lambda \underset{\sim}{x} \in K_A$ (scalar multiplication).

all remaining axioms are already satisfied in $\mathbb{F}^n$; so K_A is a vector space.

Computing $\dim(K_A)$.

let $A = \begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 3 & -3 & 1 \end{pmatrix}$; then $K_A = \{ \underset{\sim}{x} \in \mathbb{F}^4 : A\underset{\sim}{x} = \underset{\sim}{0} \}$.

to find K_A, we reduce to rref: $\begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 3 & -3 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & -2 & 2 & 0 \\ 0 & 2 & -2 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

as such, $K_A = \left\{ \begin{pmatrix} -x_4 \\ x_3 \\ x_3 \\ x_4 \end{pmatrix} \right\}$; $\boxed{\dim(K_A) = 2}$.

to obtain a basis, we take obvious choices: $x_3 = 1, x_4 = 0 \Rightarrow E_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$; $x_3 = 0, x_4 = 1 \Rightarrow E_2 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.

$\underset{\sim}{x} \in K_A \Rightarrow \underset{\sim}{x} = x_3 E_1 + x_4 E_2$; so $\{E_1, E_2\}$ span K_A and are LI ; forming a basis for K_A.

22 November 2011
Prof FEA Johnson.
Darwin LT.

$\mathbb{F}^n$ is a standard vector space.

First example of a non-standard vector space is something like this:

take field $\mathbb{F}$: take $m \times n$ matrix A over $\mathbb{F}$. consider homogeneous linear system $A\underset{\sim}{x} = \underset{\sim}{0}$, then we define $K_A = \{ \underset{\sim}{x} \in \mathbb{F}^n : A\underset{\sim}{x} = \underset{\sim}{0} \}$.

Proposition (recap) — K_A forms a vector space over $\mathbb{F}$.

Proof — let $\underset{\sim}{x}, \underset{\sim}{y} \in K_A$ i.e. $A\underset{\sim}{x} = \underset{\sim}{0}$ and $A\underset{\sim}{y} = \underset{\sim}{0} \Rightarrow A(\underset{\sim}{x} + \underset{\sim}{y}) = A\underset{\sim}{x} + A\underset{\sim}{y} = \underset{\sim}{0}$; so $\underset{\sim}{x}, \underset{\sim}{y} \in K_A \Rightarrow \underset{\sim}{x} + \underset{\sim}{y} \in K_A$.

let $\underset{\sim}{x} \in K_A,\ \lambda \in \mathbb{F}$; then $A\underset{\sim}{x} = \underset{\sim}{0}$ and $A(\lambda \underset{\sim}{x}) = \lambda(A\underset{\sim}{x}) = \lambda \cdot \underset{\sim}{0} = \underset{\sim}{0} \Rightarrow \lambda \underset{\sim}{x} \in K_A$.

finally, since $A \cdot \underset{\sim}{0} = \underset{\sim}{0}$, then $\underset{\sim}{0} \in K_A$

all other axioms are automatically satisfied since $K_A \subset \mathbb{F}^n$, and they hold in $\mathbb{F}^n \Rightarrow$ all axioms are fulfilled, K_A is a vect space.

How to find a basis for K_A:

$\boxed{\text{Ex}}$ $A = \begin{pmatrix} 1&1&1&1&1&1&1 \\ 1&0&1&0&1&0&1 \\ 1&1&0&0&0&1&0 \\ 1&2&0&1&0&2&0 \end{pmatrix}$, take $\mathbb{F} = \mathbb{Q}$; $K_A = \{ \underset{\sim}{x} \in \mathbb{Q}^7 , A\underset{\sim}{x} = \underset{\sim}{0} \}$.

$$A \longleftrightarrow \begin{pmatrix} 1&1&1&1&1&1&1 \\ 0&-1&0&-1&0&-1&0 \\ 0&0&-1&-1&-1&0&-1 \\ 0&1&-1&0&-1&1&-1 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1&0&1&0&1&0&1 \\ 0&1&0&1&0&1&0 \\ 0&0&-1&-1&-1&0&-1 \\ 0&0&-1&-1&-1&0&-1 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1&0&1&0&1&0&1 \\ 0&1&0&1&0&1&0 \\ 0&0&1&1&1&0&1 \\ 0&0&0&0&0&0&0 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1&0&0&-1&0&0&0 \\ 0&1&0&1&0&1&0 \\ 0&0&1&1&1&0&1 \\ 0&0&0&0&0&0&0 \end{pmatrix}.$$

thus, by reducing A, we see that A is row-equivalent to $\begin{pmatrix} 1&0&0&-1&0&0&0 \\ 0&1&0&1&0&1&0 \\ 0&0&1&1&1&0&1 \\ 0&0&0&0&0&0&0 \end{pmatrix}$.

also, $x_1 - x_4 = 0$, $x_2 + x_4 + x_6 = 0$ and $x_3 + x_4 + x_5 + x_7 = 0$ $\Rightarrow$ $x_1 = x_4$, $x_2 = -x_4 - x_6$, $x_3 = -x_4 - x_5 - x_7$.

$$\underset{\sim}{x} = \begin{pmatrix} x_4 \\ -x_4 - x_6 \\ -x_4 - x_5 - x_7 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} \doteq x_4 \begin{pmatrix} 1\\-1\\-1\\1\\0\\0\\0 \end{pmatrix} + x_5 \begin{pmatrix} 0\\0\\-1\\0\\1\\0\\0 \end{pmatrix} + x_6 \begin{pmatrix} 0\\-1\\0\\0\\0\\1\\0 \end{pmatrix} + x_7 \begin{pmatrix} 0\\0\\-1\\0\\0\\0\\1 \end{pmatrix} \Rightarrow \dim(K_A) = 4.$$

$$\text{basis for } K_A = \left\{ \begin{pmatrix} 1\\-1\\-1\\1\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\-1\\0\\1\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\-1\\0\\0\\0\\1\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\-1\\0\\0\\0\\1 \end{pmatrix} \right\} \therefore$$

since $\underset{\sim}{x} = x_4 E_1 + x_5 E_2 + x_6 E_3 + x_7 E_4$, so $\{E_1, E_2, E_3, E_4\}$ spans K_A.

also, $\{E_1, E_2, E_3, E_4\}$ is LI $\because$ $\sum_{i=1}^{4} \lambda_i E_i = \underset{\sim}{0} \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$.

LINEAR MAPPINGS / TRANSFORMATIONS.

$\boxed{\text{Definition}}$ Let V, W be vector spaces over a field $\mathbb{F}$.

We say that a mapping $T: V \to W$ is <u>linear</u> when

i) $T(\underset{\sim}{x} + \underset{\sim}{y}) = T(\underset{\sim}{x}) + T(\underset{\sim}{y})$ $\forall \underset{\sim}{x}, \underset{\sim}{y} \in V$,

ii) $T(\lambda \underset{\sim}{x}) = \lambda T(\underset{\sim}{x})$ $\forall \underset{\sim}{x} \in V$, $\lambda \in \mathbb{F}$.

Proposition — if $T: V \to W$ is linear, then $T(\underset{\sim}{0}) = \underset{\sim}{0}$.

Proof — $\underset{\sim}{0} = \underset{\sim}{0} + \underset{\sim}{0}$, so by property (i), $T(\underset{\sim}{0}) = T(\underset{\sim}{0} + \underset{\sim}{0}) = T(\underset{\sim}{0}) + T(\underset{\sim}{0})$.

$\therefore T(\underset{\sim}{0}) - T(\underset{\sim}{0}) = T(\underset{\sim}{0}) + T(\underset{\sim}{0}) - T(\underset{\sim}{0}) \Rightarrow 0 = T(\underset{\sim}{0})$ q.e.d.

$\boxed{\text{Ex}}$ Let $P_n(\mathbb{Q})$ be the set of polynomials of degree $\leq n$ with $\mathbb{Q}$ coefficients.

$P_n(\mathbb{Q}) = \{ a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n ; a_i \in \mathbb{Q} \}$. e.g. $P_4(\mathbb{Q}) = \{ a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 : a_i \in \mathbb{Q} \}$.

Proposition — $P_n(\mathbb{Q})$ forms a vector space over $\mathbb{Q}$.

Proof — $\underset{\sim}{a}(x) = a_0 + a_1 x + \cdots + a_n x^n$ and $\underset{\sim}{b}(x) = b_0 + b_1 x + \cdots + b_n x^n$; then

$\underset{\sim}{a}(x) + \underset{\sim}{b}(x) = (a_0 + b_0) + (a_1 + b_1) x + \cdots + (a_n + b_n) x^n$; $\underset{\sim}{a} + \underset{\sim}{b} \in P_n(\mathbb{Q})$.

likewise, for $k \in \mathbb{Q}$, $\lambda \underset{\sim}{a} = (\lambda a_0) + (\lambda a_1) x + \cdots + (\lambda a_n) x^n \in P$

zero polynomial : $0 = 0 + 0x + 0x^2 + \cdots + 0x^n \in P_n(\mathbb{Q}) \Rightarrow$ vector space, q.e.d.

Question — what is the dimension of $P_n(\mathbb{Q})$? Answer: $n+1$.

$\{1, x, x^2, \ldots, x^n\}$ forms a basis for $P_n(\mathbb{Q})$.

$\boxed{\text{Ex}}$ Let $D: P_n(\mathbb{Q}) \to P_{n-1}(\mathbb{Q})$; where $D(a(x)) = \frac{d\, a(x)}{dx}$

then $D(a+b) = D(a) + D(b)$. $\because$ $\frac{d}{dx}(a+b) = \frac{da}{dx} + \frac{db}{dx}$; likewise $\frac{d}{dx}(\lambda a) = \lambda \left(\frac{da}{dx}\right) \Rightarrow D(\lambda a) = \lambda D(a)$

hence, the function D is linear.

$\boxed{\text{Ex}}$ Let A be a $m \times n$ matrix over $\mathbb{F}$.

Consider the mapping $T_A : \mathbb{F}^n \to \mathbb{F}^m$; $T_A(\underset{\sim}{x}) = A\underset{\sim}{x}$.

then $T_A(\underset{\sim}{x} + \underset{\sim}{y}) = A(\underset{\sim}{x} + \underset{\sim}{y}) = A\underset{\sim}{x} + A\underset{\sim}{y} = T_A(\underset{\sim}{x}) + T_A(\underset{\sim}{y})$

$T_A(\lambda \underset{\sim}{x}) = A(\lambda \underset{\sim}{x}) = \lambda A \underset{\sim}{x} = \lambda T_A(\underset{\sim}{x})$;

hence T_A is linear.

We aim to show that <u>every linear map</u> behaves like the standard example.

Why are bases useful?

Proposition — suppose $\{\mathbb{E}_1, \mathbb{E}_2, \ldots, \mathbb{E}_n\}$ forms a basis for V; a vector space over $\mathbb{F}$.

then any $\underline{x} \in V$ can be expressed in a <u>unique</u> way as a linear combination $\underline{x} = \lambda_1 \mathbb{E}_1 + \lambda_2 \mathbb{E}_2 + \cdots + \lambda_n \mathbb{E}_n$ (where $\lambda_i \in \mathbb{F}$).

Proof — by the spanning property, $\underline{x}$ can be expressed in the form $\sum_{i=1}^{n} \lambda_i \mathbb{E}_i$.

the uniqueness of the expression follows as a consequence of linear independence.

suppose $\exists \mu_i \in \mathbb{F}$ s.t. $\underline{x} = \sum \lambda_i \mathbb{E}_i = \sum \mu_i \mathbb{E}_i$

since $\underline{x} - \underline{x} = 0$, $\sum \lambda_i \mathbb{E}_i - \sum \mu_i \mathbb{E}_i = \underline{0} \Rightarrow \sum (\lambda_i - \mu_i) \mathbb{E}_i = \underline{0}$.

we know that $\{\mathbb{E}_1, \mathbb{E}_2, \ldots, \mathbb{E}_n\}$ is linearly independent $\Rightarrow$ only solution is the trivial solution. hence,

$\lambda_i - \mu_i = 0 \quad \forall i \leq n$; $\lambda_i = \mu_i$ for all $i \leq n$, representation is unique, q.e.d.

this is a fundamental property of linear maps!

Proposition — let $T: V \to W$ be linear, and $\{\mathbb{E}_1, \mathbb{E}_2, \ldots, \mathbb{E}_n\}$ form a basis for V. Then T is entirely determined by the values $T(\mathbb{E}_1), T(\mathbb{E}_2), \cdots, T(\mathbb{E}_n)$.

Proof — suppose I have $T(\mathbb{E}_i) = \underline{w}_i \in W$; then we can calculate the value of $T(\underline{x})$ uniquely.

$\underline{x} = \lambda_1 \mathbb{E}_1 + \lambda_2 \mathbb{E}_2 + \cdots + \lambda_n \mathbb{E}_n$.

$T(\underline{x}) = T(\lambda_1 \mathbb{E}_1) + T(\lambda_2 \mathbb{E}_2) + \cdots + T(\lambda_n \mathbb{E}_n) = \lambda_1 T(\mathbb{E}_1) + \lambda_2 T(\mathbb{E}_2) + \cdots + \lambda_n T(\mathbb{E}_n) = \lambda_1 \underline{w}_1 + \lambda_2 \underline{w}_2 + \cdots + \lambda_n \underline{w}_n$ // q.e.d.

A linear map is determined by, and determines, a <u>matrix</u>.

let T be a linear mapping, and $T: V \to W$

we let $\{\mathbb{E}_1, \mathbb{E}_2, \ldots, \mathbb{E}_n\}$ be a basis for V, and $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m\}$ be a basis for W.

we know that T is determined by $T(\mathbb{E}_i)$; then $T(\mathbb{E}_1) \in W$ is a linear combination

← note the order of the indices … a_{mn}, not a_{nm}!

$$\left.\begin{aligned} T(\mathbb{E}_1) &= a_{11}\varepsilon_1 + a_{21}\varepsilon_2 + \cdots + a_{m1}\varepsilon_m \\ T(\mathbb{E}_2) &= a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots + a_{m2}\varepsilon_m \\ &\vdots \\ T(\mathbb{E}_n) &= a_{1n}\varepsilon_1 + a_{2n}\varepsilon_2 + \cdots + a_{mn}\varepsilon_m \end{aligned}\right\} T(\mathbb{E}_i) = \sum_{j=1}^{m} a_{ji}\,\varepsilon_j$$

$\boxed{\text{Definition}}$ Let T be linear, $T: V \to W$.

let $\mathcal{E} = \{\mathbb{E}_1, \mathbb{E}_2, \ldots, \mathbb{E}_n\}$ be a basis for V, $\Phi = \{\varphi_1, \varphi_2, \ldots, \varphi_m\}$ be a basis for W.

we obtain a matrix $M(T)_{\mathcal{E}}^{\Phi} = (a_{ji})_{1 \leq j \leq m,\, 1 \leq i \leq n}$ [$m \times n$ matrix].

defined by $T(\mathbb{E}_i) = \sum_{j=1}^{m} a_{ji}\varphi_j$

this is the matrix of T with respect to $\mathcal{E}$ on the left, Φ on the right.

$\boxed{\text{Ex}}$ Let $V = W = P_3(\mathbb{Q}) = \{a_0 + a_1 x + a_2 x^2 + a_3 x^3 : a_i \in \mathbb{Q}\}$

then $\mathcal{E} = \Phi = \{1, x, x^2, x^3\}$

take $T = D^2 + D = \frac{d^2}{dx^2} + \frac{d}{dx}$.

$$\begin{aligned} T(1) &= \tfrac{d^2}{dx^2}(1) + \tfrac{d}{dx}(1) = 0 &&= 0 + 0x + 0x^2 + 0x^3 \\ T(x) &= \tfrac{d^2}{dx^2}(x) + \tfrac{d}{dx}(x) = 1 &&= 1 + 0x + 0x^2 + 0x^3 \\ T(x^2) &= \tfrac{d}{dx^2}(x^2) + \tfrac{d}{dx}(x^2) = 2+2x &&= 2 + 2x + 0x^2 + 0x^3 \\ T(x^3) &= \tfrac{d}{dx}(x^3) + \tfrac{d}{dx}(x^3) = 6x + 3x^2 &&= 0 + 6x + 3x^2 + 0x^3 \end{aligned}$$

so $M(T)_{\mathcal{E}}^{\mathcal{E}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 0 & 6 & 3 & 0 \end{pmatrix}^T = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ //.

24 November 2011
Prof FEA Johnson.
Damin LT.

<u>Recap</u> if $T: V \to W$ is a linear map, and $\mathcal{E} = \{\mathbb{E}_1, \mathbb{E}_2, \ldots, \mathbb{E}_n\}$ form a basis of vectors for V and Φ is a basis for W.

then T is determined precisely by the values of $T(\mathbb{E}_1)$; $T(\mathbb{E}_2), \ldots, T(\mathbb{E}_n)$.

Each $T(\mathbb{E}_i)$ is uniquely a linear combination in $\{\varphi_1, \varphi_2, \ldots, \varphi_m\}$.

so we write $T(\mathbb{E}_i) = \sum_{j=1}^{m} a_{ji}\varphi_j$ (sensible convention) $= \sum_{j=1}^{m} \varphi_j a_{ji}$

alternative convention: $T(\mathbb{E}_i) = \sum_{j=1}^{m} \alpha_{ij}\varphi_j$ (be careful of element and indices) → formed by right-multiplying transformation matrix.

we define $M(T)_{\mathcal{E}}^{\Phi} = (a_{ji})_{1 \leq j \leq m,\, 1 \leq i \leq n}$: $M(T)_{\mathcal{E}}^{\Phi}$ is the matrix of T with respect to $\mathcal{E}$ on the left Φ on the right.

Composition Formula.

Let $T: V \to W$, $S: W \to X$ be two linear relations.

Let $\mathcal{E} = \{E_1, \dots, E_n\}$ be a basis for V, $\Phi = \{\varphi_1, \dots, \varphi_m\}$ be a basis for W, $\Psi = \{\psi_1, \dots, \psi_p\}$ be a basis for X.

then $S \circ T: V \to X$ is linear (check it!) $S \circ T(x) = S[T(x)]$, so

Now we have $M(T)^{\Phi}_{\mathcal{E}} \overset{(m \times n)}{=} \{(a_{ji})\ 1 \le j \le m,\ 1 \le i \le n\}$; and $M(S)^{\Psi}_{\Phi} \overset{(p \times m)}{=} \{(b_{kj})_{1 \le k \le p,\ 1 \le j \le m}\}$.

also, $M(S \circ T)^{\Psi}_{\mathcal{E}} \overset{(p \times n)}{=} \{(c_{ki})_{1 \le k \le p,\ 1 \le i \le n}\}$

$T(E_i) = \sum_{j=1}^{m} a_{ji} \varphi_j$, $S(\varphi_j) = \sum_{k=1}^{p} b_{kj} \psi_k$; and then $(S \circ T)(E_i) = \sum_{k=1}^{p} c_{ki} \psi_k$

hence, we have the relation: $\boxed{\underbrace{M(S \circ T)^{\Psi}_{\mathcal{E}}}_{p \times n} = \underbrace{M(S)^{\Psi}_{\Phi}}_{p \times m} \times \underbrace{M(T)^{\Phi}_{\mathcal{E}}}_{m \times n}}$

i.e. the matrix product corresponds exactly to composition!

Proof — $(S \circ T)(E_i) = S[T(E_i)]$, and we know $\sum_{j=1}^{m} a_{ji} \varphi_j$, so

$= S(\sum_{j=1}^{m} a_{ji} \varphi_j) = \sum_{j=1}^{m} a_{ji} S(\varphi_j)$ (because S is linear)

$= \sum_{j=1}^{m} a_{ji} \left(\sum_{k=1}^{p} b_{kj} \psi_k \right) = \sum_{j=1}^{m} \sum_{k=1}^{p} a_{ji} b_{kj} \psi_k$ observe that since $a_{ji}, b_{kj} \in \mathbb{F}$, so $a_{ji} b_{kj} = b_{kj} \cdot a_{ji}$

$= \sum_{j=1}^{m} \sum_{k=1}^{p} b_{kj} a_{ji} \psi_k = \sum_{k=1}^{p} \{ \sum_{j=1}^{m} b_{kj} a_{ji} \} \psi_k$ (interchanging order of summation) * see below

also, $(S \circ T)(E_i) = \sum_{k=1}^{p} c_{ki} \psi_k$; so by comparing the two expressions,

$c_{ki} = \sum_{j=1}^{m} b_{kj} a_{ji} = (k,i)^{th}$ entry of BA. $\Rightarrow C = BA$

$M(S \circ T)^{\Psi}_{\mathcal{E}} = M(S)^{\Psi}_{\Phi} \times M(T)^{\Phi}_{\mathcal{E}}$ // q.e.d.

Interchange of order of summation:

Suppose $\xi = (\xi_{st})_{1 \le s \le 3,\ 1 \le t \le 2}$.

$\sum_{s=1}^{3} \sum_{t=1}^{2} (\xi_{st}) = \sum_{s=1}^{3} (\xi_{s1} + \xi_{s2}) = (\xi_{11} + \xi_{12}) + (\xi_{21} + \xi_{22}) + (\xi_{31} + \xi_{32})$

$= (\xi_{11} + \xi_{21} + \xi_{31}) + (\xi_{12} + \xi_{22} + \xi_{32}) = \sum_{t=1}^{2} (\xi_{1t} + \xi_{2t} + \xi_{3t}) = \sum_{t=1}^{2} \sum_{s=1}^{3} (\xi_{st})$ // so,

$\sum_{s=1}^{3} \sum_{t=1}^{2} (\xi_{st}) = \sum_{t=1}^{2} \sum_{s=1}^{3} (\xi_{st})$ // qed.

reason for applying — (ξ_{st}) elements are in $\mathbb{F}$, so commutativity and associativity apply.

Ex (cont'd) Defining, for $P_3(\mathbb{Q}) = V$, $D: V \to V$ $D(a(x)) = \frac{Da}{Dx}$

if $\Phi = \mathcal{E} = \{1, x, x^2, x^3\}$; find $M(D)^{\mathcal{E}}_{\mathcal{E}}$, $M(D^2)^{\mathcal{E}}_{\mathcal{E}}$ and $M(D)^{\mathcal{E}}_{\mathcal{E}} M(D)^{\mathcal{E}}_{\mathcal{E}}$.

$D(1) = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3$

$D(x) = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 + 0 \cdot x^3$

$D(x^2) = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 + 0 x^3$

$D(x^3) = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2 + 0 \cdot x^3$

so $M(D)^{\mathcal{E}}_{\mathcal{E}} = \begin{pmatrix} 0&0&0&0 \\ 1&0&0&0 \\ 0&2&0&0 \\ 0&0&3&0 \end{pmatrix}^T = \begin{pmatrix} 0&1&0&0 \\ 0&0&2&0 \\ 0&0&0&3 \\ 0&0&0&0 \end{pmatrix}$ //.

$[M(D)^{\mathcal{E}}_{\mathcal{E}}]^2 = \begin{pmatrix} 0&1&0&0 \\ 0&0&2&0 \\ 0&0&0&3 \\ 0&0&0&0 \end{pmatrix} \begin{pmatrix} 0&1&0&0 \\ 0&0&2&0 \\ 0&0&0&3 \\ 0&0&0&0 \end{pmatrix} = \begin{pmatrix} 0&0&2&0 \\ 0&0&0&6 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}$ //.

Likewise,

$D^2(1) = 0$ ~~0·1 + ...~~

$D^2(x) = 0$

$D^2(x^2) = 2 \cdot 1$

$D^2(x^3) = 0 \cdot 1 + 6 \cdot x$

so $M(D^2)^{\mathcal{E}}_{\mathcal{E}} = \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 2&0&0&0 \\ 0&6&0&0 \end{pmatrix}^T = \begin{pmatrix} 0&0&2&0 \\ 0&0&0&6 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}$ // $= [M(D)^{\mathcal{E}}_{\mathcal{E}}]^2$.

Note that the order of the basis matters, i.e. permutations of basis in domain or codomain will change the order of the matrix columns.

change of basis:
we want to express each E_i as a linear combination in Φ_j where $\mathcal{E}, \Phi$ represent different bases for V.

$E_i = \sum_{j=1}^{n} a_{ji} \varphi_j$, so we have $A = (a_{ji}) = M(Id)^{\Phi}_{\mathcal{E}}$, also $\varphi_j = \sum_{i=1}^{n} b_{ij} E_i$. so we have $B = (b_{ij}) = M(Id)^{\mathcal{E}}_{\Phi}$.

29 November 2011
Prof FM to
Domin LT.

recall: if $T: V \to W$ is linear; $\mathcal{E} = \{E_1, \dots, E_n\}$ is a basis of V and $\Phi = \{\varphi_1, \varphi_2, \dots, \varphi_m\}$ is a basis for W, then

$T(E_i) = \sum_{j=1}^{m} a_{ji} \varphi_j$ and $M(T)^{\Phi}_{\mathcal{E}} = (a_{ji})_{1 \le j \le m,\ 1 \le i \le n}$.

likewise if $S: W \to U$ is also linear, and $\Psi = \{\psi_1, \psi_2, \dots, \psi_p\}$ is a basis for U; $M(S)^{\Psi}_{\Phi} = (b_{kj})_{1 \le k \le p,\ 1 \le j \le m}$

$M(S \circ T)^{\Psi}_{\mathcal{E}} = M(S)^{\Psi}_{\Phi} M(T)^{\Phi}_{\mathcal{E}}$.

1201-026

Special Case I: $W = V$ and $T: \mathrm{Id}: V \to V$; also

$\mathcal{E} = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ and $\Phi = \{\varphi_1, \ldots, \varphi_m\}$ are both bases for V.

I can express each ε_i as a linear combination in $\{\varphi_1, \ldots, \varphi_m\}$.

$$\varepsilon_i = b_{1i}\varphi_1 + b_{2i}\varphi_2 + \cdots + b_{ni}\varphi_n = \sum_{j=1}^{n} b_{ji}\varphi_j$$

and so we know that $M(\mathrm{Id})^{\Phi}_{\mathcal{E}} = (b_{ji})_{1 \le i,j \le n} = B$

likewise, we can express each φ_j as a linear combination in $\{\varepsilon_1, \ldots, \varepsilon_n\}$.

$$\varphi_j = a_{1j}\varepsilon_1 + a_{2j}\varepsilon_2 + \cdots + a_{nj}\varepsilon_n = \sum_{i=1}^{n} a_{ij}\varepsilon_i$$

and so we know that $M(\mathrm{Id})^{\mathcal{E}}_{\Phi} = (a_{ij})_{1 \le i,j \le n} = A$

then we know that $A = B^{-1}$.

i.e. $M(\mathrm{Id})^{\Phi}_{\mathcal{E}} = [M(\mathrm{Id})^{\mathcal{E}}_{\Phi}]^{-1}$.

Proof — we know that $M(\mathrm{Id})^{\mathcal{E}}_{\mathcal{E}} = I_n$; and likewise, the composition formula implies that

$$M(\mathrm{Id})^{\Phi}_{\mathcal{E}}\, M(\mathrm{Id})^{\mathcal{E}}_{\Phi} = M(\mathrm{Id})^{\mathcal{E}}_{\mathcal{E}} = I_n \Rightarrow M(\mathrm{Id})^{\Phi}_{\mathcal{E}} = [M(\mathrm{Id})^{\mathcal{E}}_{\Phi}]^{-1} \text{ q.e.d.}$$

$\boxed{\text{Ex}}$ let $V = \mathbb{Q}^3$. If $\mathcal{E} = \left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$ and $\Phi = \left\{\begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}0\\1\\3\end{pmatrix}, \begin{pmatrix}1\\2\\-1\end{pmatrix}\right\}$ are bases for V, find $M(\mathrm{Id})^{\mathcal{E}}_{\Phi}$ and $M(\mathrm{Id})^{\Phi}_{\mathcal{E}}$.

we can immediately express Φ in terms of $\mathcal{E}$.

$$\left.\begin{aligned} \varphi_1 &= 0\varepsilon_1 + 0\varepsilon_2 + 1\varepsilon_3 \\ \varphi_2 &= 0\varepsilon_1 + 1\varepsilon_2 + 3\varepsilon_3 \\ \varphi_3 &= 1\varepsilon_1 + 2\varepsilon_2 - 1\varepsilon_3 \end{aligned}\right\} \text{ so } M(\mathrm{Id})^{\mathcal{E}}_{\Phi} = \begin{pmatrix}0&0&1\\0&1&3\\1&2&-1\end{pmatrix}^T = \begin{pmatrix}0&0&1\\0&1&2\\1&3&-1\end{pmatrix}$$

then $M(\mathrm{Id})^{\Phi}_{\mathcal{E}} = \begin{pmatrix}0&0&1\\0&1&2\\1&3&-1\end{pmatrix}^{-1} = \begin{pmatrix}7&-3&1\\-2&1&0\\1&0&0\end{pmatrix}$

$$\left(\begin{array}{ccc|ccc}0&0&1&1&0&0\\0&1&2&0&1&0\\1&3&-1&0&0&1\end{array}\right) \longleftrightarrow \left(\begin{array}{ccc|ccc}1&3&-1&0&0&1\\0&1&2&0&1&0\\0&0&1&1&0&0\end{array}\right) \longleftrightarrow \left(\begin{array}{ccc|ccc}1&3&0&1&0&1\\0&1&0&-2&1&0\\0&0&1&1&0&0\end{array}\right) \longleftrightarrow \left(\begin{array}{ccc|ccc}1&0&0&7&-3&1\\0&1&0&-2&1&0\\0&0&1&1&0&0\end{array}\right).$$

or, from first principles,

$$\varepsilon_1 = \begin{pmatrix}1\\0\\0\end{pmatrix} = a_{11}\begin{pmatrix}0\\0\\1\end{pmatrix} + a_{21}\begin{pmatrix}0\\1\\3\end{pmatrix} + a_{31}\begin{pmatrix}1\\2\\-1\end{pmatrix} \Rightarrow \begin{pmatrix}0&0&1\\0&1&2\\1&3&-1\end{pmatrix}\begin{pmatrix}a_{11}\\a_{21}\\a_{31}\end{pmatrix} = \begin{pmatrix}1\\0\\0\end{pmatrix} \Rightarrow a_{11} = 7,\ a_{21} = -2,\ a_{31} = 1$$

$$\varepsilon_2 = \begin{pmatrix}0\\1\\0\end{pmatrix} = a_{12}\begin{pmatrix}0\\0\\1\end{pmatrix} + a_{22}\begin{pmatrix}0\\1\\3\end{pmatrix} + a_{32}\begin{pmatrix}1\\2\\-1\end{pmatrix} \quad \cdots \quad a_{12} = -3,\ a_{22} = 1,\ a_{32} = 0$$

$$\varepsilon_3 = \begin{pmatrix}0\\0\\1\end{pmatrix} = a_{13}\begin{pmatrix}0\\0\\1\end{pmatrix} + a_{23}\begin{pmatrix}0\\1\\3\end{pmatrix} + a_{33}\begin{pmatrix}1\\2\\-1\end{pmatrix} \quad \cdots \quad a_{13} = 1,\ a_{23} = 0,\ a_{33} = 0$$

$M(\mathrm{Id})^{\Phi}_{\mathcal{E}} = \begin{pmatrix}7&-3&1\\-2&1&0\\1&0&0\end{pmatrix}$ q.e.d.

Special Case II: (change of basis formula). Here we have $T: V \to V$; where once again, we define $\mathcal{E}$ and Φ as both bases for V.

Similar case as above, except that T may not be the identity. (more general).

we compare matrices $M(T)^{\mathcal{E}}_{\mathcal{E}}$ with $M(T)^{\Phi}_{\Phi}$; then $\underbrace{M(T)^{\Phi}_{\Phi}}_{B} = M(\mathrm{Id})^{\Phi}_{\mathcal{E}}\, \underbrace{M(T)^{\mathcal{E}}_{\mathcal{E}}}_{A}\, M(\mathrm{Id})^{\mathcal{E}}_{\Phi}$.

Proof — composition formula: $\mathrm{LHS} = M(\mathrm{Id})^{\Phi}_{\mathcal{E}} M(T)^{\mathcal{E}}_{\mathcal{E}} M(\mathrm{Id})^{\mathcal{E}}_{\Phi} = M(\mathrm{Id} \circ T \circ \mathrm{Id})^{\Phi}_{\Phi} = M(T)^{\Phi}_{\Phi}$ q.e.d.

Nomenclature: $M(\mathrm{Id})^{\Phi}_{\mathcal{E}}$, $M(\mathrm{Id})^{\mathcal{E}}_{\Phi}$ are matrices of the "change of basis". (or transition matrices).

letting $P = M(\mathrm{Id})^{\Phi}_{\mathcal{E}}$; $\boxed{B = PAP^{-1}}$ alternative notation.

$\boxed{\text{Ex}}$ Let $V = \mathbb{R}^2$. $T: \mathbb{R}^2 \to \mathbb{R}^2$, $T\underline{x} = \begin{pmatrix}2x_1 + x_2\\x_1 + 2x_2\end{pmatrix}$. Let $\Phi = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\right\}$. Find $M(T)^{\Phi}_{\Phi}$.

If $\mathcal{E} = \left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$, $M(T)^{\mathcal{E}}_{\mathcal{E}} = \begin{pmatrix}2&1\\1&2\end{pmatrix}$.

$M(\mathrm{Id})^{\mathcal{E}}_{\Phi}$ can be found by letting φ_j be a linear combination in $\mathcal{E}$; $= \begin{pmatrix}1&1\\1&-1\end{pmatrix}$

$M(\mathrm{Id})^{\Phi}_{\mathcal{E}} = [M(\mathrm{Id})]^{\mathcal{E}}_{\Phi} = -\frac{1}{2}\begin{pmatrix}-1&-1\\-1&1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix} = \begin{pmatrix}\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&-\frac{1}{2}\end{pmatrix}$.

so $M(T)^{\Phi}_{\Phi} = \frac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}1&1\\1&-1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}3&3\\1&-1\end{pmatrix}\begin{pmatrix}1&1\\1&-1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}6&0\\0&2\end{pmatrix} = \begin{pmatrix}3&0\\0&1\end{pmatrix}$

$\boxed{\text{Ex}}$ Let $T: \mathbb{Q}^3 \to \mathbb{Q}^3$ be a linear map, $T(\underline{x}) = \begin{pmatrix}2x_2 + 2x_3\\3x_1 + x_2 - 2x_3\\x_1 - x_2 + 2x_3\end{pmatrix}$

If $\mathcal{E}$ is the standard basis, $M(T)^{\mathcal{E}}_{\mathcal{E}} = \begin{pmatrix}0&2&2\\3&1&-2\\1&-1&2\end{pmatrix}$. Take $\Phi = \left\{\begin{pmatrix}1\\-1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\-1\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right\}$. Calculate $M(T)^{\Phi}_{\Phi}$.

then $M(\mathrm{Id})^{\mathcal{E}}_{\Phi} = \begin{pmatrix}1&0&0\\-1&1&0\\1&-1&1\end{pmatrix} \Rightarrow M(\mathrm{Id})^{\Phi}_{\mathcal{E}} = \begin{pmatrix}1&0&0\\-1&1&0\\1&-1&1\end{pmatrix}^{-1} = \begin{pmatrix}1&0&0\\1&1&0\\0&1&1\end{pmatrix}$.

$$\because \left(\begin{array}{ccc|ccc}1&0&0&1&0&0\\-1&1&0&0&1&0\\1&-1&1&0&0&1\end{array}\right) \longleftrightarrow \left(\begin{array}{ccc|ccc}1&0&0&1&0&0\\0&1&0&1&1&0\\0&-1&1&-1&0&1\end{array}\right) \longleftrightarrow \left(\begin{array}{ccc|ccc}1&0&0&1&0&0\\0&1&0&1&1&0\\0&0&1&0&1&1\end{array}\right)$$

then $M(T)^{\Phi}_{\Phi} = M(\mathrm{Id})^{\Phi}_{\mathcal{E}}\, M(T)^{\mathcal{E}}_{\mathcal{E}}\, M(\mathrm{Id})^{\mathcal{E}}_{\Phi}$

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

$$= \begin{pmatrix}0&2&2\\3&3&0\\4&0&0\end{pmatrix}\begin{pmatrix}1&0&0\\1&1&0\\0&1&1\end{pmatrix} = \begin{pmatrix}0&0&2\\0&3&0\\4&0&0\end{pmatrix}$$

Special case III: Let $A = (a_{ji})_{1 \le j \le m,\ 1 \le i \le n}$ be an $m \times n$ matrix over $\mathbb{F}$.

Consider the linear map $T_A : \mathbb{F}^n \to \mathbb{F}^m$; $T_A(\underline{x}) = A\underline{x}$.

Take $\varepsilon = \{\underline{e}_1, \dots, \underline{e}_n\}$ be the standard basis for $\mathbb{F}^n$; $\epsilon = \{\underline{\epsilon}_1, \dots, \underline{\epsilon}_m\}$ be the standard basis for $\mathbb{F}^m$.

What is matrix $M(T_A)^{\epsilon}_{\varepsilon}$?

calculate $T_A(\underline{e}_1) = A\underline{e}_1 = \begin{pmatrix} a_{11} & \cdots\cdots & a_{1n} \\ a_{21} & \cdots\cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots\cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_{11}\epsilon_1 + a_{21}\epsilon_2 + \cdots a_{m1}\epsilon_m = \sum_{j=1}^{m} a_{ji}\epsilon_j$. ~~etc~~

so 1^{st} column of $M(T_A)^{\epsilon}_{\varepsilon}$ = 1^{st} column of A.

so $\boxed{M(T_A)^{\epsilon}_{\varepsilon} = A}$.

Corollary – matrix multiplication is associative.

Proof – Let A be an $m \times n$ matrix over $\mathbb{F}$; B be $n \times p$ and C be $p \times q$.

We want to show $(AB)C = A(BC)$. Then we define:

$T_A : \mathbb{F}^n \to \mathbb{F}^m$ $(T_A(\underline{x}) = A\underline{x})$; $T_B : \mathbb{F}^p \to \mathbb{F}^n$ $(T_B(\underline{y}) = B\underline{y})$; $T_C : \mathbb{F}^q \to \mathbb{F}^p$ $(T_C(\underline{z}) = C\underline{z})$.

$$\begin{tikzcd} \mathbb{F}^q \arrow[r, "T_C"] \arrow[rr, bend left] & \mathbb{F}^p \arrow[r, "T_B"] \arrow[rr, bend right] & \mathbb{F}^n \arrow[r, "T_A"] & \mathbb{F}^m \end{tikzcd}$$

then $T_A \circ (T_B \circ T_C) = (T_A \circ T_B) \circ T_C$ since composition is associative.

$M(T_A \circ (T_B \circ T_C))^{std}_{std} = M(T_A)^{std}_{std}\, M(T_B \circ T_C)^{std}_{std} = A\left[M(T_B)^{std}_{std}\, M(T_C)^{std}_{std}\right] = A(BC)$.

$M((T_A \circ T_B) \circ T_C)^{std}_{std} = M(T_A \circ T_B)^{std}_{std}\, M(T_C)^{std}_{std} = \left[M(T_A)^{std}_{std}\, M(T_B)^{std}_{std}\right] C = (AB)C$.

thus, $A(BC) = (AB)C$, q.e.d.

1 December 2011
Prof FEA Johnson.
Darwin LT.

PRINCIPLES OF ABSTRACT ALGEBRA.

Let V be a vector space over field $\mathbb{F}$; $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n \in V$.

We say that $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ is LI when $\sum_{i=1}^{n} \lambda_i \underline{v}_i = \underline{0} \iff \forall i,\ \lambda_i = 0$.

$\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ spans V when, given any $\underline{x} \in V$, we can write $\underline{x} = \lambda_1\underline{v}_1 + \cdots + \lambda_n\underline{v}_n$ for some $\lambda_1, \dots, \lambda_n \in \mathbb{F}$.

$\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ forms a basis for V when both $\{\underline{v}_1, \dots, \underline{v}_n\}$ is LI and spans V.

Main result: BASIS THEOREM.

Theorem BASIS THEOREM – proof for later.

If V is a non-zero vector space then

i) V has a basis, and

ii) any two bases for V have the same number of elements.

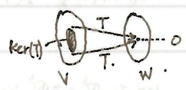
Definition $\dim(V)$ = no. of elements in a basis.

hence, by convention, $\dim(0) = 0$.

KERNEL-RANK THEOREM.

Suppose $T : V \to W$ is linear. We associate to T two additional vector spaces.

- kernel of T: $\ker(T) = \{\underline{v} \in V : T(\underline{v}) = \underline{0}\}$
- image of T: $\mathrm{Im}(T) = \{\underline{w} \in W : \text{for some } \underline{v} \in V,\ T(\underline{v}) = \underline{w}\}$



i.e. $\mathrm{Im}(T)$ consists of all elements in W which are hit by the mapping T.

Proof that $\ker(T)$ and $\mathrm{Im}(T)$ are vector spaces; and also the relation $\boxed{\dim[\ker(T)] + \dim[\mathrm{Im}(T)] = \dim(V)}$

Definition Suppose V is a vector space over $\mathbb{F}$ and $U \subset V$, we say that U is a vector subspace of V when

i) $\underline{0} \in U$, ii) $\underline{x}, \underline{y} \in U$ then $\underline{x} + \underline{y} \in U$ (closure over addition), iii) $\underline{x} \in U$ and $\lambda \in \mathbb{F}$ then $\lambda\underline{x} \in U$.

since all other axioms are already satisfied in parent subset V, U is itself also a vector space.

$\Rightarrow$ a vector subspace is itself a vector space.

$0 \in \ker(T)$ $\because$ $T(\underline{0}) = \underline{0}$ and $\underline{0} \in V$. If $\underline{x}, \underline{y} \in \ker(T)$, then $T(\underline{x}) = T(\underline{y}) = \underline{0} \Rightarrow T(\underline{x} + \underline{y}) = T(\underline{x}) + T(\underline{y}) = \underline{0}$, so $\underline{x} + \underline{y} \in \ker(T)$.

1201-028

If $\underline{x} \in Ker(T)$ and $\lambda \in \mathbb{F}$, then $T(\lambda \underline{x}) = \lambda T(\underline{x}) = \lambda \cdot 0 = 0$, so $\lambda \underline{x} \in Ker(T)$. All three conditions are fulfilled $\Rightarrow Ker(T)$ is a vector subspace of $V \Rightarrow Ker(T)$ is a vector space, q.e.d.

also, since W is a vector space, $\underline{0} \in Im(T)$ since $T(\underline{0}) = \underline{0}$. If $\underline{w}_1, \underline{w}_2 \in Im(T)$, then $\underline{w}_1 = T(\underline{v}_1)$ and $\underline{w}_2 = T(\underline{v}_2) \Rightarrow T(\underline{v}_1 + \underline{v}_2) = T(\underline{v}_1) + T(\underline{v}_2) = \underline{w}_1 + \underline{w}_2 \in Im(T)$.

If $w \in Im(T)$ and $\lambda \in \mathbb{F}$, then $\underline{w} = T(\underline{v})$ so $T(\lambda \underline{v}) = \lambda T(\underline{v}) = \lambda \underline{w}$, and $\lambda \underline{w} \in Im(T)$, q.e.d.

To prove the kernel-rank theorem, we consider the generic case where both $Ker(T) \neq 0$, $Im(T) \neq 0$.

We use the basis theorem to construct a basis $\{\underline{e}_1, \dots, \underline{e}_k\}$ for $ker(T)$, $\dim[ker(T)] = k$.

We also construct a basis $\{\varphi_1, \dots, \varphi_m\}$ for $Im(T)$, $\dim[Im(T)] = m$.

$$\left.\begin{array}{ll} \text{we choose} & \underline{e}_{k+1} \in V, \text{ then } T(\underline{e}_{k+1}) = \varphi_1 \\ & \underline{e}_{k+2} \in V, \text{ then } T(\underline{e}_{k+2}) = \varphi_2 \\ & \vdots \\ & \underline{e}_{k+m} \in V, \text{ then } T(\underline{e}_{k+m}) = \varphi_m \end{array}\right\} \begin{array}{l} \text{we claim } \{\underline{e}_1, \dots, \underline{e}_{k+m}\} \text{ is a basis for } V, \text{ i.e. need to show.} \\ \text{a) } \{\underline{e}_1, \dots, \underline{e}_{k+m}\} \text{ is LI; and} \\ \text{b) } \{\underline{e}_1, \dots, \underline{e}_{k+m}\} \text{ spans } V. \end{array}$$

a) suppose that the vectors have coefficients $\lambda_1, \lambda_2, \dots, \lambda_{k+m}$ s.t.

$\lambda_1 \underline{e}_1 + \dots + \lambda_k \underline{e}_k + \lambda_{k+1} \underline{e}_{k+1} + \dots + \lambda_{k+m} \underline{e}_{k+m} = 0$, then we must show that each $\lambda_i = 0$.

applying T, $T\left(\sum_{i=1}^{k+m} \lambda_i \underline{e}_i\right) = \sum_{i=1}^{k+m} \lambda_i T(\underline{e}_i) = T(\underline{0}) = \underline{0}$.

so $\sum_{i=1}^{k+m} \lambda_i T(\underline{e}_i) = 0$; but since $\{\underline{e}_1, \dots, \underline{e}_k\} \in ker(T)$, $T(\underline{e}_i) = \underline{0}$ for $1 \leq i \leq k$.

also, $T(\underline{e}_{k+i}) = \varphi_i$ for $1 \leq i \leq m$.

$\sum_{i=k+1}^{k+m} \lambda_i T(\underline{e}_i) = \underline{0} \Rightarrow \sum_{i=1}^{m} \lambda_{k+i} \varphi_i = \underline{0}$; but since $\{\varphi_1, \dots, \varphi_n\}$ is a basis for $Im(T)$, it is LI;

so $\lambda_{k+1} = \lambda_{k+2} = \dots = \lambda_{k+m} = 0$; so now $\sum_{i=1}^{k} \lambda_i \underline{e}_i = \underline{0}$, but $\{\underline{e}_1, \dots, \underline{e}_k\}$ form a basis for $ker(T)$ and is LI.

hence, $\forall i$, $1 \leq i \leq k+m$; $\lambda_i = 0 \Rightarrow$ set is LI, q.e.d.

b) let $\underline{v} \in V$. we must produce $\lambda_1, \dots, \lambda_{k+m} \in \mathbb{F}$ s.t. $\underline{v} = \sum_{i=1}^{k+m} \lambda_i \underline{e}_i$

applying T to $\underline{v}$, then $T(\underline{v}) \in Im(T)$

we write $T(\underline{v}) = \mu_1 \varphi_1 + \dots + \mu_m \varphi_m$ $(\mu_1, \dots; \mu_m \in \mathbb{F})$ because $\{\varphi_1, \dots, \varphi_m\}$ spans $Im(T)$.

consider $\underline{v}' \in V$ defined by $\underline{v}' = \mu_1 \underline{e}_{k+1} + \dots + \mu_m \underline{e}_{k+m}$. Applying T again, $T(v') = \mu_1 \varphi_1 + \dots + \mu_m \varphi_m$.

so $T(\underline{v}) = T(\underline{v}')$ and hence $T(\underline{v} - \underline{v}') = \underline{0}$ i.e. $\underline{v} - \underline{v}' \in Ker(T)$.

we write $\underline{v} - \underline{v}' = \lambda_1 \underline{e}_1 + \dots + \lambda_k \underline{e}_k$, so $\underline{v} = \lambda_1 \underline{e}_1 + \dots + \lambda_k \underline{e}_k + \mu_1 \underline{e}_{k+1} + \dots \mu_m \underline{e}_{k+m}$.

so $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_{k+m}\}$ spans V, q.e.d.

6 December 2011.
Prof FEA Johnson.
Darwin LT.

we have proven the kernel-rank-theorem for the generic case: i.e. $ker(T) \neq 0$, $Im(T) \neq 0$.

Now, we consider special cases –

case 1: $ker(T) = 0$.

we want $\dim(V) = \dim(Im\,T)$; then $\underline{e}_1, \dots, \underline{e}_m$ maps basis $\varphi_1, \dots, \varphi_m$ for all $Im(T)$.

$T(v - v') = \underline{0} \Rightarrow v - v' \in ker(T) = 0$.

so $v = v' = x_1 \underline{e}_1 + \dots + x_m \underline{e}_m \Rightarrow \dim V = m = \dim(Im\,T)$.

case 2: $Im(T) = 0$.

Then $T \equiv 0$ so $Ker(T) = V$, so $\dim V = \dim ker(T)$.

case 3: $Ker(T) = Im(T) = 0 \Rightarrow$ trivial.

How does this relate to what we already know?

Standard case: A is an $m \times n$ matrix over $\mathbb{F}$.

$T_A: \mathbb{F}^n \to \mathbb{F}^m$ is linear and $T_A(\underline{x}) = A\underline{x}$ (matrix product).

Here, $\dim ker(T_A) + \dim Im(T_A) = n = \dim \mathbb{F}^n$.

In this case, $Ker(T_A) = \{\underline{x} \in \mathbb{F}^n ; A\underline{x} = 0\}$ i.e. the <u>solution set</u> (in former notation, K_A).

Then, how about $Im(T_A)$?

$$Im(T_A) = \{\underline{y} \in \mathbb{F}^m : \underline{y} = A\underline{x} \text{ for some } x \in \mathbb{F}^n\}.$$

$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$; then first column of A is $A_{*1} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}$, likewise $A_{*2} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}$; $A_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$.

then since $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, then $\boxed{A\underline{x} = x_1 A_{*1} + x_2 A_{*2} + \dots x_n A_{*n}}$.

$\boxed{\text{Corollary}}$ $\underline{y} \in Im(T_A) \Leftrightarrow \underline{y}$ is a linear combination of columns in A i.e. $\underline{y} = x_1 A_{*1} + x_2 A_{*2} + \dots + x_n A_{*n}$.

1201-029.

$\boxed{\text{Definition}}$ The columnspace of $m \times n$ matrix A is the set of linear combinations in columns of A

i.e. $\lambda_1 A_{*1} + \lambda_2 A_{*2} + \cdots + \lambda_n A_{*n}$

As such, we see that $\operatorname{Im}(T_A)$ = column space of A.

Computing $\dim \operatorname{Im}(T_A)$ and $\dim \operatorname{Ker}(T_A)$.

- When A is in reduced row echlon form, it is simple.

$\boxed{\text{Ex}}$ $A = \begin{array}{c} \begin{pmatrix} 1 & 7 & 0 & 2 & 3 & 0 & 1 \\ 0 & 0 & 1 & -1 & 4 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ \begin{matrix} * & & * & & & * & \end{matrix} \end{array}$ $*$ note the columns with the leading 1s.

Everything in the columnspace can be expressed as a linear combination of these.

so $\{A_{*1}, A_{*3}, A_{*6}\}$ forms a natural basis for the columnspace $\Rightarrow$ a basis for $\operatorname{Im}(T_A)$

e.g. $A_{*7} = A_{*1} + 2A_{*3} - A_{*6}$. $\Rightarrow \dim \operatorname{Im}(T_A) = 3$

so, in a reduced row-echlon matrix, $\dim \operatorname{Im}(T_A)$ = no. of dependent variables = no. of non-zero rows = no. of leading "1"s.

then also, $\dim \operatorname{Ker}(T_A)$ = no. of dependent variables = $n - \dim \operatorname{Im}(T_A)$ (as per our knowledge of the kernel-rank theorem).

i.e. # dependent variables + # independent variables = $\dim \operatorname{Im}(T_A) + \dim \operatorname{Ker}(T_A) = n$.

- When A is not in reduced form... in general,

A is an $m \times n$ matrix over $\mathbb{F}$.

We reduce A to row echlon form A' by left-multiplying by some invertible matrix P; i.e. $A' = PA \Rightarrow A = P^{-1}A'$.

this gives us a linear map: $\binom{\text{column space}}{A} \longrightarrow \binom{\text{column space}}{A'}$. and likewise; $\binom{\text{column space}}{A'} \longrightarrow \binom{\text{column space}}{A}$.

$\underset{\sim}{c} \longrightarrow P\underset{\sim}{c}$ $\qquad$ $\underset{\sim}{y} \longrightarrow P^{-1}\underset{\sim}{y}$.

Let $\varphi_1, \ldots \varphi_t$ be a basis for (col space A'); then we claim that $P^{-1}\varphi_1, \ldots, P^{-1}\varphi_t$ is a basis for (col space A).

likewise, if $\psi_1, \psi_2, \ldots \psi_q$ is a basis for (col space A), then $P\psi_1, \ldots, P\psi_q$ is a basis for (col space A')

i.e. $t = q$; and dim (col space A) = dim (col space A') = $t = q$.

Suppose $\lambda_1 P^{-1}\varphi_1 + \cdots + \lambda_t P^{-1}\varphi_t = 0$

$P^{-1}(\lambda_1\varphi_1 + \cdots + \lambda_t\varphi_t) = 0$. $\Rightarrow$ $PP^{-1}(\lambda_1\varphi_1 + \cdots + \lambda_t\varphi_t) = P \cdot 0 = 0$

so $\lambda_1\varphi_1 + \cdots \lambda_t\varphi_t = 0$; but these are LI $\Rightarrow$ $\lambda_1 = \lambda_2 = \cdots = \lambda_t = 0$. $\Rightarrow$ $\{P^{-1}\varphi_1, \ldots, P^{-1}\varphi_t\}$ is LI.

Let $y \in$ col space A.

$Py \in \cdots A'$. write $Py = \lambda_1\varphi_1 + \cdots + \lambda_t\varphi_t$, so $y = \lambda_1 P^{-1}\varphi_1 + \cdots + \lambda_t P^{-1}\varphi_t$.

so $\{P^{-1}\varphi_1, \ldots, P^{-1}\varphi_t\}$ spans the column space of A.

so if dim (col space of A') = t, then (col space of A) has a basis with t elements.

so dim (col space A) = dim (col space A') = $q = t$.

$\boxed{\text{Ex}}$ for $\mathbb{F} = \mathbb{Q}$, we note that

$A = \begin{pmatrix} 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -2 & 1 & -1 \end{pmatrix}$, and $T_A : \mathbb{Q}^6 \to \mathbb{Q}^4$, $T_A(\underset{\sim}{x}) = A\underset{\sim}{x}$. Find a basis for $\ker(T_A)$, and for $\operatorname{Im}(T_A)$

we reduce A to row echlon form (left multiplying transformation matrices).

$$A = \begin{pmatrix} 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -2 & 1 & -1 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1 & -1 & 1 & -1 & 1 & -1 \\ 0 & 2 & 0 & 0 & -2 & 0 \\ 0 & 2 & 0 & 2 & -2 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 1 & 0 & 1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix} \longleftrightarrow \begin{array}{c} \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{matrix} \end{array}.$$

so we know that $\dim \ker(T_A) = 3$, and $\dim \operatorname{Im} T_A = 3$.

we know that $\{A'_{*1}, A'_{*2}, A'_{*4}\}$ form a basis for col. space of A' (reduced A) $\Rightarrow$ the corresponding columns form a basis for col. space of A.

$\Rightarrow \{A_{*1}, A_{*2}, A_{*4}\} = \left\{ \begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\\1\\-1\end{pmatrix}, \begin{pmatrix}-1\\-1\\1\\-2\end{pmatrix} \right\}$ form a basis for col. space of A' $\Rightarrow$ basis for $\operatorname{Im}(T_A)$.

General solution for $A\underset{\sim}{x} = \underset{\sim}{0}$ $\equiv$ general solution for $A'\underset{\sim}{x} = \underset{\sim}{0}$.

general solution is $\begin{pmatrix} -x_3 + x_6 \\ x_5 \\ x_3 \\ 0 \\ x_5 \\ x_6 \end{pmatrix}$ $\Rightarrow$ $\underset{\sim}{x} = x_3\begin{pmatrix}-1\\0\\1\\0\\0\\0\end{pmatrix} + x_5\begin{pmatrix}0\\1\\0\\0\\1\\0\end{pmatrix} + x_6\begin{pmatrix}1\\0\\0\\0\\0\\1\end{pmatrix}$. $\Rightarrow$ $\left\{ \begin{pmatrix}-1\\0\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\\0\\1\end{pmatrix} \right\}$ form a basis for $\operatorname{Ker}(T_A)$.

BASIS THEOREM.

We start by considering a simplified version of the EXCHANGE LEMMA

$\boxed{\text{Theorem}}$ Let V be a vector space over field $\mathbb{F}$. Let $\{f_1, \ldots, f_n\}$ be a spanning set for V. Let $\underset{\sim}{v} \in V$ ($\underset{\sim}{v} \neq \underset{\sim}{0}$).

Suppose $\underset{\sim}{v} = \lambda_1 f_1 + \cdots + \lambda_{r-1} f_{r-1} + \lambda_r f_r + \cdots + \lambda_n f_n$ where $\lambda_r \neq 0$, then $\{f_1, f_2, \ldots, f_{r-1}, \underset{\sim}{v}, f_{r+1}, \ldots, f_n\}$ still spans V.

Proof — since $\underset{\sim}{v} \in V$, $\underset{\sim}{v} \neq 0$, then $\underset{\sim}{v} = \lambda_r \varphi_r + \sum_{t \neq r} \lambda_t \varphi_t$ $\boxed{\lambda_r \neq 0}$

so $\varphi_r = \left(\frac{1}{\lambda_r}\right) \underset{\sim}{v} + \sum_{t \neq r} \frac{-\lambda_t}{\lambda_r} \varphi_t$ ($\because$ $\mathbb{F}$ is a field, λ_r^{-1} exists)

we claim that $\{\varphi_1, \ldots\ldots, \varphi_{r-1} \vdots \underset{\sim}{v} \vdots \varphi_{r+1}, \ldots, \varphi_n\}$ spans V. (let the set be Φ'; as compared to Φ).

so if we let $\underset{\sim}{x} \in V$, then we need to be able to express $\underset{\sim}{x}$ as a linear combination in Φ'

we also know that we can express $\underset{\sim}{x}$ s.t. it is a linear combination in Φ.

so $\underset{\sim}{x} = \mu_r \varphi_r + \sum_{t \neq r} \mu_t \varphi_t$, and by substitution,

$$\underset{\sim}{x} = \mu_r\left[\left(\frac{1}{\lambda_r}\right)\underset{\sim}{v} + \sum_{t \neq r} \frac{-\lambda_t}{\lambda_r} \varphi_t\right] + \sum_{t \neq r} \mu_t \varphi_t = \left(\frac{\mu_r}{\lambda_r}\right) \underset{\sim}{v} + \sum_{t \neq r} \left(\mu_t - \frac{\mu_r}{\lambda_r} \lambda_t\right) \varphi_t.$$

thus, we have seen that $\underset{\sim}{x}$ can indeed be expressed as a linear combination in $\Phi' \Rightarrow \Phi'$ spans V// q.e.d.

this can help us understand the full exchange lemma.

Before that, we note that if $\{w_1, \ldots, w_k\}$ is a set of vectors, if any $w_i = \underset{\sim}{0}$, then set is not linearly independent.

because if $\underset{\sim}{w}_r = \underset{\sim}{0}$, we can write $\underset{\sim}{0} = 0 \cdot \underset{\sim}{w}_1 + 0 \cdot \underset{\sim}{w}_2 + \cdots\cdots + 0 \cdot \underset{\sim}{w}_{r-1} + \mu \cdot \underset{\sim}{w}_r + \cdots + 0 \cdot \underset{\sim}{w}_k$ for $\mu \neq 0$. (not all coefficients must be 0).

$\boxed{\text{Theorem}}$ FULL EXCHANGE LEMMA. (by STEINITZ)

let V be a vector space, and $\{\varphi_1, \ldots, \varphi_n\}$ is a spanning set. Suppose $\{v_1, \ldots, v_k\}$ is a LI set in V, then

(i) $k \leq n$ and (ii) there exists a spanning set $\{\psi_1, \ldots, \psi_n\}$ for V such that $\psi_i = v_i$ for $i \leq k$, and $\psi_t \in \{\varphi_1, \ldots, \varphi_n\}$ for $n \geq t > k$.

Proof — by induction on k.

we've already shown the case for $k=1$ (simplified exchange lemma).

so suppose we have proven it for $k-1$; that means we have constructed a spanning set $\varphi_1', \ldots, \varphi_n'$ such that.

$\varphi_1' = v_1$, $\varphi_2' = v_2$, $\ldots$, $\varphi_{k-1}' = v_{k-1}$, and beyond that, for $n \geq t > k-1$, $\varphi_t' \in \{\varphi_1, \ldots, \varphi_n\}$.

we express v_k as a linear combination in $\varphi_1', \varphi_2', \ldots, \varphi_n'$

$$v_k = \sum_{t=1}^{m} \lambda_t \varphi_t' = \sum_{t=1}^{k} \lambda_t \varphi_t' + \sum_{t \geq k}^{n} \lambda_t \varphi_t'$$

since $\{v_1, v_2, \ldots, v_k\}$ is LI, $v_k \neq 0$, so some $\lambda_t \neq 0$. we claim that $\lambda_t \neq 0$ for some $t \geq k$.

(otherwise $v_k = \sum_{t=1}^{k-1} \lambda_t v_t$, which contradicts LI of $v_1, \ldots, v_k$).

choose r, $k \leq r \leq n$ s.t. $\lambda_r \neq 0$, by the simplified lemma,

$$\{v_1, \ldots, v_{k-1}, v_k \vdots \varphi_{k+1}', \varphi_{k+2}', \ldots, \varphi_n'\} = \underbrace{\{v_1, \ldots, v_{k-1} \vdots \varphi_{k+1}', \ldots \vdots v_k \vdots \ldots\ldots \varphi_n'\}}_{\text{swapping positions.}} \text{ still spans V.}$$

hence, by re-indexing,

$$\{\underset{\psi_1}{\overset{\shortparallel}{v_1}}, \ldots, v_{k-1}, \underset{\psi_k}{\overset{\shortparallel}{v_k}} \vdots \underset{\psi_{k+1}}{\overset{\shortparallel}{\varphi_{k+1}'}} \ldots \varphi_{r-1}' \vdots \varphi_{r+1}' \ldots \underset{\psi_k}{\overset{\shortparallel}{\varphi_n'}}\}.$$ // q.e.d.

8 December 2011
Prof FEA Johnson.
Darwin LT.

BASIS THEOREM.

$\boxed{\text{Theorem}}$ let V be a <u>nonzero</u> vector space over field $\mathbb{F}$; then

(i) (existence) V has at least one basis.

(ii) (uniqueness) any two bases for V have the same number of elements.

Proof of (ii) — suppose $\{\epsilon_1, \epsilon_2, \ldots, \epsilon_m\}$ is a basis for V, and also $\{\varphi_1, \ldots, \varphi_n\}$ is another basis for V.

NTP: $m=n$

Applying the exchange lemma, we see that $\{\epsilon_1, \ldots, \epsilon_m\}$ is LI, and $\{\varphi_1, \ldots, \varphi_n\}$ spans V $\Rightarrow m \leq n$.

and also, we see that $\{\varphi_1, \ldots, \varphi_n\}$ is LI, and $\{\epsilon_1, \ldots, \epsilon_m\}$ spans V $\Rightarrow n \leq m$.

thus, $m \leq n \leq m \Rightarrow m=n$// q.e.d.

Proof of (i) — V has at least one spanning set, namely V itself.

let X be a <u>minimal spanning set</u> (i.e. a spanning set with the property that it has no proper subsets which also span).

we claim that X is LI, so X is a basis. Begin by assuming otherwise, proof by contradiction

i.e. X is not LI, so we must find a relation $\lambda_1 \underset{\sim}{w}_1 + \lambda_2 \underset{\sim}{w}_2 + \cdots + \lambda_p \underset{\sim}{w}_p = \underset{\sim}{0}$ for some $\lambda_i \neq 0$, all $\underset{\sim}{w}_i \in X$

$\lambda_p \underset{\sim}{w}_p = -(\lambda_1 w_1 + \cdots + \lambda_{p-1} \underset{\sim}{w}_{p-1}) \Rightarrow \underset{\sim}{w}_p = \sum_{i=1}^{p-1} \left(-\frac{\lambda_i}{\lambda_p}\right) \underset{\sim}{w}_i$. Define $X' = X \setminus \{w_p\}$. As X spans, X' also spans.

But $X' \underset{\text{proper}}{\subset} X \Rightarrow$ contradiction of minimality. $\Rightarrow$ X is LI $\Rightarrow$ X is a basis// q.e.d.

$\blacksquare$

note: the existence proof we have covered is simple if we assume that V has a finite spanning set.

However, for some vector spaces, they have infinite spanning sets (will be covered in "FUNCTIONAL ANALYSIS").

13 December 2011
Prof FEA Johnson
Dermin LT.

PERMUTATIONS (cont'd).

A permutation on n letters is a bijective mapping $\sigma : \{1,2,\ldots,n\} \longrightarrow \{1,2,\ldots,n\}$.

We have shown that

i) Any permutation σ is a product $\sigma = c_1 \ldots c_m$.

where each c_i is a cyclic permutation, and c_i, c_j are disjoint for $i \neq j$. In particular $c_i c_j = c_j c_i \quad \forall i,j$.

$\boxed{\text{Definition}}$ The order of σ = the least $N \geq 1$ such that $\sigma^N = Id$.

(by convention, $\sigma^0 = Id$).

Because $c_i c_j = c_j c_i$, then $\sigma^k = c_1^k c_2^k \ldots c_m^k$, each c_i^k is a cycle.

In order for $\sigma^N = Id$, we need each $c_i^N = Id$. the order of c_i = length of c_i.

Proposition — $\text{ord}(\sigma) = \text{lcm}(\text{length } c_i)$, where $\sigma = c_1 \ldots c_m$ (product of disjoint cycles).

$\boxed{\text{Ex}}$ $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 4 & 12 & 1 & 3 & 13 & 5 & 7 & 6 & 2 & 10 & 9 & 11 & 8 \end{pmatrix}$

then $\sigma = (1,4,3)(2,12,11,9)(5,13,8,6)(7)(10)$, $c_4 = c_5 = Id$

$\sigma^k = c_1^k c_2^k c_3^k$ $\quad c_1^k = Id$ only where k is a multiple of 3; likewise $c_2^k - 4$ and $c_3^k - 4 \Rightarrow$ so.

in order to get $\sigma^k = Id$, we need k to be a multiple of both 3 and 4.

$\text{ord}(\sigma) = \text{lcm}(3,4) = 12$ //

We have also shown that

i) a cycle of length m is a product of $m-1$ transpositions.

ii) a transposition is a product of an odd number of adjacent transpositions.

iii) when length $(c) = m$ is odd, then c is a product of an even number of adjacent transpositions.

when length $(c) = m$ is even, then c is a product of an odd number of adjacent transpositions.

$\boxed{\text{Definition}}$ sign of σ, $\text{sign}(\sigma)$ is defined such that:

- if σ is a product of an even number of adjacent transpositions, $\text{sign}(\sigma) = 1$.
- if σ is a product of an odd number of adjacent transpositions, $\text{sign}(\sigma) = -1$.

This only makes sense if we can show that:

$\boxed{\text{Theorem}}$ It is impossible to write a permutation σ as a product of both an even and odd number of adjacent transpositions.

Laplace's Formula for the sign of a permutation.

$\sigma : \{1,\ldots,n\} \longrightarrow \{1,\ldots,n\}$.

we define $\boxed{\pi(\sigma) = \dfrac{\prod_{1 \leq i < j \leq n} \sigma(j) - \sigma(i)}{\prod_{1 \leq i < j \leq n} (j-i)}}$ then there are $\frac{n(n-1)}{2}$ terms in each product.

then we will show that $\text{sign}(\sigma) = \pi(\sigma)$: this is a theoretical tool, not a practical method.

Define $L(\sigma) = \prod_{1 \leq i < j \leq n} \sigma(j) - \sigma(i)$ $\quad \left[\text{also that } \pi(\sigma) = \frac{L(\sigma)}{L(Id)}\right]$.

Proposition — if τ is an adjacent transposition, then $L(\sigma\tau) = -L(\sigma)$;

for some permutation σ.

1201-032.

Proof — we fix the transposition τ as $\tau = (p, p+1)$.

we distinguish several sets (seven in total):

$\mathcal{F}(1) = \{(i,j) : 1 \le i < j < p\}$.

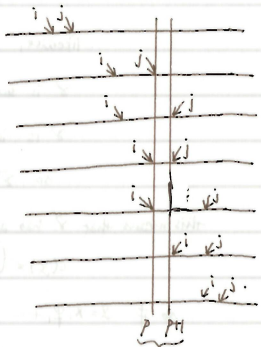
$\mathcal{F}(2) = \{(i,j) : 1 \le i < j = p\}$

$\mathcal{F}(3) = \{(i,j) : 1 \le i < p < j = p+1\}$

$\mathcal{F}(4) = \{(i,j) : i = p,\ j = p+1\}$

$\mathcal{F}(5) = \{(i,j) : i = p < p+1 < j\}$

$\mathcal{F}(6) = \{(i,j) : i = p+1 < j\}$

$\mathcal{F}(7) = \{(i,j) : p+1 < i < j\}$

only places where a swap occurs.

Define $L_r(\sigma) = \prod_{(i,j) \in \mathcal{F}(r)} (\sigma(j) - \sigma(i))$, so $L(\sigma) = \prod_{r=1}^{7} L_r(\sigma)$.

now we check.

$L_1(\sigma\tau) = L_1(\sigma)$, $L_2(\sigma\tau) = L_3(\sigma)$, $L_3(\sigma\tau) = L_2(\sigma)$, $L_4(\sigma\tau) = -L_4(\sigma)$.

$L_5(\sigma\tau) = L_6(\sigma)$, $L_6(\sigma\tau) = L_5(\sigma)$, $L_7(\sigma\tau) = L_7(\sigma)$.

so, $\prod_{r=1}^{7} L_r(\sigma\tau) = -\prod_{r=1}^{7} L_r(\sigma)$, and thus $L(\sigma\tau) = -L(\sigma)$, q.e.d.

Corollary — let σ be a permutation, $\tau_1, \dots, \tau_k$ be adjacent transpositions.

then $L(\sigma\tau_1 \cdots \tau_k) = (-1)^k L(\sigma)$.

Proof — Each time we perform a τ operation, we change the sign by -1.

Corollary — If $\tau_1, \dots, \tau_k$ are adjacent transpositions, then $\Pi(\tau_1 \cdots \tau_k) = (-1)^k$.

Proof — take $\sigma = \mathrm{Id}$, then $L(\tau_1 \cdots \tau_k) = (-1)^k L(\mathrm{Id})$, so.

$\frac{L(\tau_1 \cdots \tau_k)}{L(\mathrm{Id})} = (-1)^k$, i.e. $\Pi(\tau_1 \cdots \tau_k) = (-1)^k$, q.e.d.

Corollary — $\Pi(\tau_1 \cdots \tau_k) = \begin{cases} +1 & k \text{ even} \\ -1 & k \text{ odd.} \end{cases}$

Corollary — sign of permutation is well-defined, and $\mathrm{sign}(\sigma) = \Pi(\sigma)$.

Corollary — $\Pi(\rho\sigma) = \Pi(\rho)\Pi(\sigma)$.

Proof — write $\rho = \tau_1 \cdots \tau_k$, $\sigma = \tau_{k+1} \cdots \tau_{k+m}$.

then $\Pi(\rho\sigma) = (-1)^{k+m} = (-1)^k(-1)^m = \Pi(\rho)\Pi(\sigma)$, q.e.d.

Corollary — $\mathrm{sign}(\rho\sigma) = \mathrm{sign}(\rho)\,\mathrm{sign}(\sigma)$

so if $\sigma = c_1 \dots c_m$, c_i are disjoint cycles,

$\mathrm{sign}(\sigma) = \mathrm{sign}(c_1) \cdots \mathrm{sign}(c_m) = (-1)^{\mathrm{length}(c_1)-1} \cdots (-1)^{\mathrm{length}(c_m)-1}$.

Ex How to compute the sign — in practice. Let $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 6 & 3 & 2 & 1 & 4 & 9 & 14 & 10 & 5 & 13 & 7 & 11 & 8 & 12 \end{pmatrix}$. Find $\mathrm{sign}(\sigma)$.

$\sigma = (1, 6, 9, 5, 4)(2, 3)(7, 14, 12, 11)(8, 10, 13)$

$\mathrm{sign}(\sigma) = \prod_{r=1}^{4} \mathrm{sign}(c_r) = \prod_{r=1}^{4} (-1)^{\mathrm{length}(c_r)-1} = (-1)^{5-1}(-1)^{2-1}(-1)^{4-1}(-1)^{3-1} = (-1)^4(-1)^1(-1)^3(-1)^2 = (+1)(-1)(-1)(+1) = +1$.

$\mathrm{order}(\sigma) = \mathrm{lcm}(5, 2, 4, 3) = 60$.

DIMENSION.

What exactly do we mean by the "dimension" of a vector space?

by definition, $\dim(V)$ = number of elements in a basis.

suppose $\dim(V) = m \ge 1$ and let $\{\varphi_1, \dots, \varphi_m\}$ be a basis.

we know that $\mathbb{F}^m$ also has dimension m, and $\mathbb{F}^m$ has a standard basis $\{e_1, \dots, e_m\}$, where $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ etc.

Any vector $\underline{x} \in V$ can be expressed uniquely in the form $\underline{x} = x_1\varphi_1 + \dots + x_n\varphi_m$ $x_i \in \mathbb{F}$

likewise any vector $\underline{x} \in \mathbb{F}^m$ can be expressed uniquely as $\underline{x} = x_1 e_1 + \dots + x_m e_m$

Consider the mapping $\gamma : \mathbb{F}^n \to V$, where $\gamma\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\varphi_1 + x_2\varphi_2 + \dots + x_m\varphi_m$.

then we have:

Proposition — γ is a linear map, and γ is bijective.

Proof — $\gamma(\underset{\sim}{x}+\underset{\sim}{y}) = \gamma\begin{pmatrix} x_1+y_1 \\ \vdots \\ x_m+y_m \end{pmatrix} = (x_1+y_1)\varphi_1 + \cdots + (x_m+y_m)\varphi_m = x_1\varphi_1 + \cdots + x_m\varphi_m + y_1\varphi_1 + \cdots + y_m\varphi_m$

$= \gamma(\underset{\sim}{x}) + \gamma(\underset{\sim}{y})$.

likewise, $\gamma(\lambda \underset{\sim}{y}) = \lambda\gamma(\underset{\sim}{y})$ // linear. q.e.d.

γ is surjective because $\{\varphi_1, \dots, \varphi_m\}$ spans V. i.e. $\forall \underset{\sim}{v} \in V$ we can write $\underset{\sim}{v} = x_1\varphi_1 + \cdots + x_m\varphi_m = \gamma(\underset{\sim}{x})$.

γ is injective $\because \{\varphi_1, \dots, \varphi_m\}$ is LI. $\gamma(\underset{\sim}{x}) = \gamma(\underset{\sim}{y}) \Rightarrow (x_1-y_1)\varphi_1 + \cdots + (x_m-y_m)\varphi_m = \underset{\sim}{0} \Rightarrow x_i = y_i \because$ LI of $\Phi \Rightarrow \underset{\sim}{x} = \underset{\sim}{y}$.

so γ is bijective // q.e.d.

this means that γ has an inverse map $c: V \longrightarrow \mathbb{F}^m$, called the coordinate map.

$c(\underset{\sim}{v}) = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \iff \underset{\sim}{v} = x_1\varphi_1 + x_2\varphi_2 + \cdots\cdots + x_m\varphi_m$

so if $\underset{\sim}{v} = x_1\varphi_1 + \cdots + x_m\varphi_m$, then $(x_1, \dots, x_m)$ are the "coordinates" of v with reference to $\{\varphi_1, \dots, \varphi_m\}$.

coordinates were a mainstay of 19th century mathematics, until D. Hilbert's innovation of geometric notation: using $\underset{\sim}{x} = x_1\underset{\sim}{e}_1 + \dots + x_m\underset{\sim}{e}_m$.

so, we deduce that a vector space of dimension m behaves exactly like $\mathbb{F}^m$.

Not every V is an $\mathbb{F}^m$ but it does behave like $\mathbb{F}^m$.

$\boxed{\text{Ex}}$ consider $V = \{\underset{\sim}{x} \in \mathbb{F}^3 : x_1 + x_2 + x_3 = 0\}$.

V has dimension 2.

quick method to demonstrate: write $T: \mathbb{F}^3 \to \mathbb{F}$, where $T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 + x_2 + x_3$.

T is linear.

$\underset{\sim}{x} \in V \iff T(\underset{\sim}{x}) = 0$.

T is surjective. $T\begin{pmatrix} \lambda \\ 0 \\ 0 \end{pmatrix} = \lambda$ for any λ. $V = \ker(T)$. $\Rightarrow \dim\ker(T) + \dim \operatorname{Im}(T) = 3$, so $\dim(V) + 1 = 3$ so $\dim(V) = 2$.

standard method gives basis for V.

matrix of T w.r.t. standard bases. $\underset{x_1\ x_2\ x_3}{(1, 1, 1)} \Rightarrow x_1 = -x_2 - x_3$, $\varphi_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$, $\varphi_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

any $\underset{\sim}{x} \in V$ has form $\underset{\sim}{x} = x_2\varphi_1 + x_3\varphi_2$; so $V \neq \mathbb{F}^2$.

However V behaves exactly like $\mathbb{F}^2$.

$\boxed{\text{Definition}}$ Let V, W be vector spaces over $\mathbb{F}$. We say that V and W are <u>isomorphic</u> (notation: $V \cong W$) if $\exists$ bijective linear map $T: V \to W$

$\boxed{\text{Theorem}}$ $V \cong \mathbb{F}^m$ (isomorphic) $\iff \dim(V) = m$

$V \cong W \iff \dim(V) = \dim(W)$.

if $T: V \to W$ is linear bijective, then T is said to be an <u>isomorphism</u>.

Proposition — if $T: V \to W$ is linear bijective, then $T^{-1}: W \to V$ is also linear bijective.

Proof — bijection of T^{-1} automatically shown by existence of T^{-1} for T; so we only need to show linearity.

we consider $T^{-1}(\underset{\sim}{w}_1 + \underset{\sim}{w}_2) - T^{-1}(\underset{\sim}{w}_1) - T^{-1}(\underset{\sim}{w}_2)$. Then if we apply T, since T is linear,

$T[T^{-1}(\underset{\sim}{w}_1 + \underset{\sim}{w}_2) - T^{-1}(\underset{\sim}{w}_1) - T^{-1}(\underset{\sim}{w}_2)] = T \circ T^{-1}(\underset{\sim}{w}_1 + \underset{\sim}{w}_2) - T \circ T^{-1}(\underset{\sim}{w}_1) - T \circ T^{-1}(\underset{\sim}{w}_2) = \underset{\sim}{w}_1 + \underset{\sim}{w}_2 - \underset{\sim}{w}_1 - \underset{\sim}{w}_2 = \underset{\sim}{0}$.

but T is injective, so $T^{-1}(\underset{\sim}{w}_1 + \underset{\sim}{w}_2) - T^{-1}(\underset{\sim}{w}_1) - T^{-1}(\underset{\sim}{w}_2) = 0$, and $T^{-1}(\underset{\sim}{w}_1 + \underset{\sim}{w}_2) = T^{-1}(\underset{\sim}{w}_1) + T^{-1}(\underset{\sim}{w}_2)$.

likewise, we consider. $T[T^{-1}(\lambda\underset{\sim}{w}) - \lambda T^{-1}(\underset{\sim}{w})] = 0$, so $T^{-1}(\lambda\underset{\sim}{w}) = \lambda T^{-1}(\underset{\sim}{w})$ by analogous argument // q.e.d.

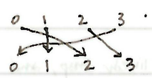
other properties of isomorphism —

- $V \cong W \iff W \cong V$.
- $\underbrace{U \cong V}_{T}$ and $\underbrace{V \cong W}_{S} \Rightarrow \underbrace{U \cong W}_{S \circ T}$

FINITENESS.

The concepts of finite and infinite sets.

Imagine if $f: \{0,1,2,3\} \to \{0,1,2,3\}$ is an injection. e.g.



note that if f is injective, it is automatically surjective.

1201-034.

On the other hand, consider $\mathbb{N} = \{1, 2, \dots\}$ where $g: \mathbb{N} \to \mathbb{N}$, $g(n) = n+1$

where g is injective here, it still is not surjective $\because 1 \notin \mathrm{Im}(g)$.

$\boxed{\text{Definition}}$ A set A is called <u>finite</u> when every injective map $f: A \to A$ is also surjective.

For instance, $\{1, 2, \dots, n-1, n\}$ is finite, but $\mathbb{N}$ is infinite.

Question: To what extent is $\mathbb{N}$ a "typical" infinite set?

$\boxed{\text{Definition}}$ An infinite set A is called countable when there exists a bijective mapping $f: \mathbb{N} \to A$. (where $0 \in \mathbb{N}$).

For instance, $\mathbb{N}$ is countable, take $f = \mathrm{Id}$.

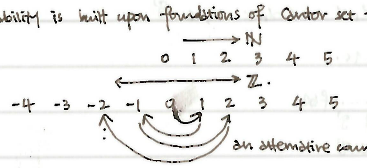
$\mathbb{Z}^+$ is countable, since $\mathbb{Z}^+ = \{n \in \mathbb{N} : 1 \leq n\} = \{1, 2, \dots, n, n+1\}$.

However, this leads us to Galileo's paradox.

$\mathbb{Z}^+$ is countable because $f: \mathbb{N} \setminus \{0\} \to \mathbb{N}$, $f(n) = n+1$ is bijective; $f^{-1}(m) = m-1$ $(m \geq 1)$.

This implies that one can take a proper subset of an infinite set, which is still infinite. Which infinity is larger?

Our notion of countability is built upon foundations of Cantor set theory.



an alternative counting principle.

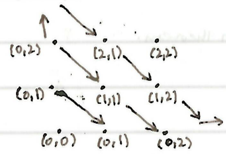
Is $\mathbb{Z}$ countable? Yes.

We define $g: \mathbb{N} \to \mathbb{Z}$.

$g(2n) = n$, $g(2n+1) = -(n+1)$. [two cases]

Then g is bijective.

Take $\mathbb{N} \times \mathbb{N}$; is that countable? Yes.



Here, we approach the problem the other way round.

Consider $\Psi: \mathbb{N} \times \mathbb{N} \to \mathbb{N}$.

$$\Psi(m,n) = \frac{(m+n)(m+n+1)}{2} + n.$$

For instance, $\Psi(0,0) = 0$, $\Psi(1,0) = 1$, $\Psi(0,1) = 2$, $\Psi(2,0) = 3 \dots\dots$

We know that Ψ is bijective, so $\Psi^{-1}: \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ is also bijective $\Rightarrow$ $\mathbb{N} \times \mathbb{N}$ is countable.

Corollary — If A, B are countable infinite sets, then $A \times B$ is countable.

Proof — take bijections $f: \mathbb{N} \to A$, $g: \mathbb{N} \to B$.

Define $\Phi: \mathbb{N} \to A \times B$ by $\Phi = [f, g]\Psi^{-1}$

$$\begin{tikzcd} \mathbb{N} \arrow[rr, "\Phi"] \arrow[dr, "\Psi"'] & & A \times B \\ & \mathbb{N} \times \mathbb{N} \arrow[ur, "f \circ g"'] & \end{tikzcd}$$

Corollary — If $A_1, A_2, \dots, A_m$ are countable, then $A_1 \times A_2 \times \dots \times A_m$ is also countable.

Proof — think of $B = A_1 \times A_2 \times \dots \times A_{m-1}$, then $A_1 \times A_2 \times \dots \times A_{m-1} \times A_m = B \times A_m$.

by induction, B is countable, so hence $B \times A_m$ is countable.

Think of 'mathematical alphabet' as being a countable set, which includes:

$\mathcal{A} = \{+, -, \cdot, \div, \mathbb{N}, a, b, \dots, z, a', b', \dots, z', \dots\}$, then $\mathcal{A}$ is a countable set.

This implies that a mathematical expression of length n is an element in $\mathcal{A}_n$, where $\mathcal{A}_n = \underbrace{\mathcal{A} \times \mathcal{A} \times \dots \times \mathcal{A}}_{(n \text{ times})}$.

Each $\mathcal{A}_n$ is countable, so, the totality of mathematical expressions is $\bigcup_{n \geq 1} \mathcal{A}_n$.

Proposition — $\bigcup_{n \geq 1} \mathcal{A}_n$ is countable.

Proof — For each n, let $\varphi_n: \mathbb{N} \to \mathcal{A}_n$ be bijective.

Get $c: \underset{(m,n)}{\mathbb{N} \times \mathbb{N}} \longrightarrow \underset{\varphi_n(m)}{\bigcup \mathcal{A}_n}$, so $c \circ \Psi^{-1}: \mathbb{N} \to \bigcup_{n \geq 1} \mathcal{A}_n$ is bijective, q.e.d.

$\boxed{\text{Theorem}}$ LÖWENHEIM-SKOLEM THEOREM: The set of all possible mathematical expressions is countable.

In Analysis, one is required to believe the following statement.

<u>Uncountability of $\mathbb{R}$</u> (Cantor) : the set $\mathbb{R}$ is not countable.

Proof by analogy, <u>Cantor's diagonal trick</u>:

let $\mathscr{P}$ be infinite subsets of $\mathbb{N}$, Cantor showed that $\mathscr{P}$ is not countable.

we write each element $\alpha \in \mathscr{P}$ as an increasing sequence, i.e.

$$\alpha = (\alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_n < \alpha_{n+1} < \cdots)$$

Suppose $\mathscr{P}$ is ~~invertible~~ countable

$$\alpha^0 = (\alpha_0^0 < \alpha_1^0 < \cdots < \alpha_n^0 < \cdots$$

$$\alpha^1 = (\alpha_0^1 < \alpha_1^1 < \cdots < \alpha_n^1 < \cdots$$

$$\vdots$$

$$\alpha^n = (\alpha_0^n < \alpha_1^n < \cdots < \alpha_n^n < \cdots$$

Then we define $\beta \in \mathscr{P}$ by $\beta_n = \sum_{0 \le i,j \le n} \alpha_j^i + n$; then $\beta_n \in \mathbb{N}$.

we see then, clearly, that

β is not in the list because $\alpha_r^r < \beta_r$ for any r.

which contradicts countability assumption, q.e.d.

Then... we see that $\mathbb{R}$ is at least as big as $\mathscr{P}$.

let $\alpha = (\alpha_0 < \alpha_1 < \cdots < \alpha_n < \cdots) \in \mathscr{P}$

Write down decimal with 0, 1s.

$$f(\alpha) = f(\alpha)_0\, f(\alpha)_1 \ldots\ldots f(\alpha)_n$$

$$f(\alpha)_r = \begin{cases} 0 & \text{if } r \notin \mathscr{P} \\ 1 & \text{if } r \in \mathscr{P} \end{cases}$$

so $\mathbb{R}$ is not countable.

This leads, in turn, to an argument with the Löwenheim-Skolem theorem.

END OF SYLLABUS.