HANDOUT 7

MATRICES

- Introduction
  - consistency and number of solutions of 2x2 systems
  - definition of a matrix
  - matrix arithmetic
    - * addition and subtraction
    - * multiplication by a scalar
    - * multiplying matrices
- Solution of equations
  - Cramer's rule
  - Laplace expansion of determinants
  - Classification of systems I.

HANDOUT 8

CONTENTS

- Classification of systems of linear equations
  II. The Rank of a matrix
- Linear Independence of
  - equations
  - vectors
  - functions

HANDOUT 9

CONTENTS

- Properties of determinants
- Special matrices (I)
- Inverse of a matrix
  - Formal method
  - Row reduction method
- Special matrices (II)
- Applications of matrices
  - Solution of simultaneous linear equations
  - Rotation of a reference frame
  - General linear transformations
  - Systems theory
- Eigenvalues and Eigenvectors

HANDOUT 7

MATRICES

- Introduction
  - consistency and number of solutions of 2×2 systems
  - definition of a matrix
  - matrix arithmetic
    - addition and subtraction
    - multiplication by a scalar
    - multiplying matrices
- Solution of equations
  - Cramer's rule
  - Laplace expansion of determinants
  - Classification of systems I.

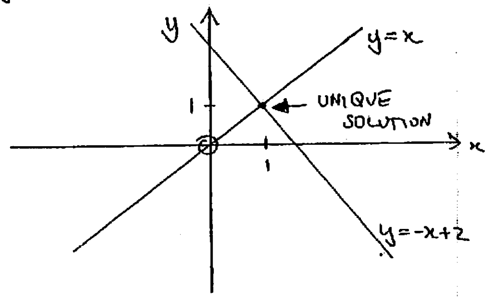
MATRICES

Introduction

The solution of simultaneous equations is a problem that appears regularly in everyday life and throughout science and engineering. Systems involving, say, just two linear equations are easy to solve and this can be done either graphically or by manipulation of the equations

Ex
$$\begin{aligned} x+y&=2 \\ x-y&=0 \end{aligned}$$
We want to find x and y (the unknowns). Adding the two equations gives $2x=2$. Thus, $x=1$ and substitution of this value in either equation yields $y=1$.

... graphically, we have the lines $y=-x+2$ and $y=x$

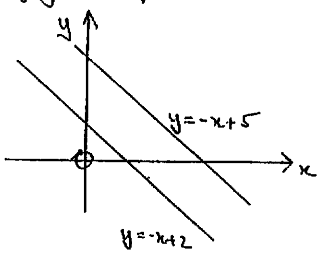


i.e. we have 2 equations and 2 unknowns, giving a unique solution.

Ex
$$\begin{aligned} x+y&=2 \\ x+y&=5 \end{aligned}$$
2 equations and 2 unknowns again. Subtract the top equation from the bottom one to find
$$\begin{aligned} x+y&=2 \\ 0.x+0.y&=3 \end{aligned} \quad \left(\text{coefficients are } \begin{matrix} 1 & 1 \\ 0 & 0 \end{matrix}\right)$$

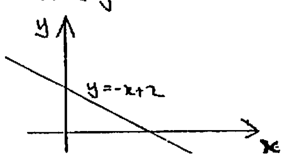
Woops! That gives $0=3$. Something's wrong. The equations are "INCONSISTENT" i.e. we can't have x and y simultaneously satisfying both equations

Graphically ...
parallel lines
⇒ they never cross
⇒ no solution



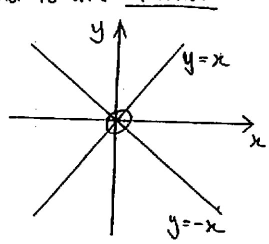
Ex
$$\begin{aligned} x+y&=2 \\ 2x+2y&=4 \end{aligned}$$
2 equations and 2 unknowns again? No! The second equation is precisely just twice the first. They are basically the same equation i.e. $x+y=2$ and we really only have 1 equation with 2 unknowns. This is just a line ...

→ we have an infinite number of solutions
i.e. all the points on the line.



Ex A "homogeneous" system i.e. the right hand side has zeroes.
$$\begin{aligned} x+y&=0 \\ x-y&=0 \end{aligned}$$
2 equations and 2 unknowns
- one isn't a multiple of the other, so we do have 2 independent equations
- they are not inconsistent
  $y=-x$ and $y=x$ are not parallel lines

BUT, the only solution we can find is the "trivial" or "null" solution $x=y=0$



Ex Another "homogeneous" system.
$$\begin{aligned} x+y&=0 \\ 2x+2y&=0 \end{aligned}$$
This time, one equation is a multiple of the other. They are consistent – they must be, they are the same equation

Take 2 times the first equation away from the second ...

... to give
$$\begin{aligned} x+y&=0 \\ 0.x+0.y&=0 \end{aligned}$$
i.e. coefficients $\begin{matrix} 1 & 1 \\ 0 & 0 \end{matrix}$

The second just says $0=0$ i.e. it doesn't give any information and we are just left with the first equation (194)

$x+y=0$ i.e. $y=-x$.

So, with one equation and two unknowns, the homogeneous system gives an infinite number of solutions.

This is all easy in the case of 2 equations, but what do we do if there are 3, 4, 5 or more equations?

We need systematic ways of determining ...

- whether solutions exist
- how many exist
- finding them.

In other words, we need to

- write the equations in a way that they can be analysed (MATRICES)
- work out how to answer the above questions (MATRIX THEORY)

So, what is a matrix? (195)

Consider the system of equations
$$\begin{aligned} x + 2y &= 5 \\ 3x - y &= 8 \end{aligned}$$

The coefficients of x and y "$\begin{matrix} 1 & 2 \\ 3 & -1 \end{matrix}$" form a "matrix".

i.e. a rectangular array of "elements" giving a table of values. To keep things tidy, we put some square or round brackets around the table and give it a name [NOT a value, it's a table of values, just a name for now]

e.g. $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$ this is a "2×2" matrix.

We know that the solutions depend on both this and the right hand side of the equations, so let's define another matrix $\begin{bmatrix} 5 \\ 8 \end{bmatrix}$ ... this is a "2×1" matrix ... it's a (short) column ... it's like a column vector.

We could call this $B = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$, but because it also represents a vector I might write it as $\underset{\sim}{b} \equiv \begin{bmatrix} 5 \\ 8 \end{bmatrix}$.

196

Similarly, let's write the solution to the system of equations (the unknowns) in terms of a matrix $X = \begin{bmatrix} x \\ y \end{bmatrix}$.

Again, for these type of problems, it will usually be just a single column, like a vector.

Thus, I may write $\underset{\sim}{x} \equiv \begin{bmatrix} x \\ y \end{bmatrix}$ to emphasize that this matrix is a single column.

What other types of matrix could we have?

The coefficients don't need to be numeric. They could be constants $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where a, b, c, d are real or complex.

They could be expressions, say polynomials $\begin{bmatrix} \alpha & \alpha+1 \\ \alpha^2+2 & \beta^7+1 \end{bmatrix}$.

They could even be other matrices!

Before we can write our system of equations in terms of matrices, we need to know how to manipulate matrices (add, subtract, multiply, etc.) ... the rules of the game →

197

Matrix arithmetic

(a) Matrix addition

Let $A = \begin{pmatrix} 2 & 1 & 4 \\ -3 & 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 3 & -5 & 1 \\ 2 & 1 & 3 \end{pmatrix}$

then $A+B = \begin{pmatrix} 2+3 & 1-5 & 4+1 \\ -3+2 & 0+1 & 2+3 \end{pmatrix}$

$= \begin{pmatrix} 5 & -4 & 5 \\ -1 & 1 & 5 \end{pmatrix}$

Easy! But what if they are not the same "order"?

In the above, A and B have 2 rows and 3 columns i.e. they are "2×3", read as "2 by 3" matrices. If they were of different order then one simply could not add them.

(b) Matrix subtraction

$A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$, $B = \begin{pmatrix} g & h & i \\ j & k & \ell \end{pmatrix}$

$A - B = \begin{pmatrix} a-g & b-h & c-i \\ d-j & e-k & f-\ell \end{pmatrix}$

Easy! But, again, they must be of the same order!

198

(c) Multiplying a matrix by a number (a "scalar")

Let's say "λ times A" where $A = \begin{pmatrix} 2 & 1 & 4 \\ -3 & 0 & 2 \end{pmatrix}$.

$\lambda A = \begin{pmatrix} 2\lambda & \lambda & 4\lambda \\ -3\lambda & 0 & 2\lambda \end{pmatrix}$

i.e. we multiply every single element by λ.

e.g. $\lambda = 2$ gives $\lambda A = \begin{pmatrix} 4 & 2 & 8 \\ -6 & 0 & 4 \end{pmatrix}$.

(d) Multiplying one matrix by another (more involved!)

- when can we do it?
- how do we do it?

ex $A = \begin{pmatrix} 2 & 1 & 4 \\ -3 & 0 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 5 \\ 2 & -1 \\ 4 & 2 \end{pmatrix}$

a 2×3 matrix   a 3×2 matrix

Then, $AB = \begin{bmatrix} (2)(3)+(1)(2)+(4)(4) & (2)(5)+(1)(-1)+(4)(2) \\ (-3)(3)+(0)(2)+(2)(4) & (-3)(5)+(0)(-1)+(2)(2) \end{bmatrix}$

i.e. $AB = \begin{pmatrix} 24 & 17 \\ -1 & -11 \end{pmatrix}$ ...... a 2×2 matrix

199

What have we done?

- To get row 1, column 1 element of AB (i.e. 24) we multiplied corresponding elements of ... row 1 of A and column 1 of B

... it was like a scalar product of vectors

i.e. $(2, 1, 4)$ times $\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$ gave $(2)(3) + (1)(2) + (4)(4) = 24$

Compare with with vectors $\underset{\sim}{u} = (u_1, u_2, u_3)$
$\underset{\sim}{v} = (v_1, v_2, v_3)$

then $\underset{\sim}{u} \cdot \underset{\sim}{v} = u_1v_1 + u_2v_2 + u_3v_3$ = "a number" / "a scalar"

- To get row 2, column 1 element of AB, we did the same operation on row 2 and column 1 of matrices A and B, respectively
- Similarly, row 1 of A and column 2 of B gives the element in row 1 and column 2 of AB.
- and similarly for the row 2, column 2 element of AB

But these operations like dot products are only possible (200)
if the number of elements in the rows of A EQUALS the number of elements in the columns of B.

Now, the number of elements in a row = the number of columns of the matrix

and the number of elements in a column = the number of rows of the matrix

Finally(!), we can multiply A and B if the number of columns of A = the number of rows of B

e.g. $A = \begin{pmatrix} 2 & 1 & 4 \\ -3 & 0 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 5 \\ 2 & -1 \\ 4 & 2 \end{pmatrix}$

$2 \times \underbrace{3 \qquad 3}\times 2$

These need to be equal if we want to multiply the matrices. The matrices are said to be "CONFORMABLE".

The order (or "size") of AB is then given by the outer numbers

i.e in this case, AB is a 2X2 matrix.

i.e. the result of multiplying A and B is a 2X2 matrix.

(201)

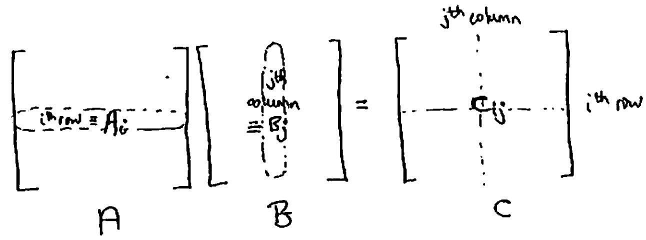
General notation

For a matrix A (as above) one usually denotes the elements as a_{ij} where $\boxed{\begin{array}{l} i \text{ is the ROW subscript} \\ j \text{ is the COLUMN subscript} \end{array}}$

e.g. $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$

So, in general, to multiply two matrices • THINK SCALAR PRODUCT and • VISUALISE THIS

to get the ij^{th} element of $AB = C$



where scalar $c_{ij} \equiv \underset{\sim}{A_i} \cdot \underset{\sim}{B_j} = [a_{i1} \; a_{i2} \; \cdots \; a_{in}] \cdot \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix}$

$= a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj}$

(202)

••• and this is only possible when A and B are conformable

i.e. $\begin{array}{ccc} A & B & = C \\ n \times m & m \times p & n \times p \end{array}$

Some multiplication properties

- $A(BC) = (AB)C$ associative
- $A(B+C) = AB + AC$ } distributive
- $(B+C)A = BA + CA$ } distributive
- $AB \neq BA$ non-commutation

Another example of "row dot column":

$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$

$2 \times \underbrace{2 \qquad 2}_{\text{conformable}} \times 2$

$AB = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$ ..... a 2X2 matrix results

(203)

Solutions of equations

Having defined matrix multiplication, we can now express a set of simultaneous linear equations in matrix form.

e.g. $ax + by = e$
$cx + dy = f$

is $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} e \\ f \end{pmatrix}$

$(2\times2) \; (2\times1) \to (2\times1)$

I would tend to write this as $\boxed{A\underset{\sim}{x} = \underset{\sim}{b}}$

where $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ≡ coefficient matrix

$\underset{\sim}{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ ≡ solution vector

$\underset{\sim}{b} = \begin{pmatrix} e \\ f \end{pmatrix}$ ≡ "right hand side" determining whether the system is homogeneous $\underset{\sim}{b} = \underset{\sim}{0}$ or inhomogeneous $\underset{\sim}{b} \neq \underset{\sim}{0}$.

Let's solve $\begin{matrix} ax + by = e \\ cx + dy = f \end{matrix}$ and see what the result is.

multiply the first equation by d and the second by b to get

$$(ad)x + (bd)y = ed$$
$$(cb)x + (db)y = fb$$

subtract $\quad (ad-bc)x = (ed-fb)$

$$\Rightarrow \boxed{x = \frac{ed-fb}{ad-bc}}$$

Substituting this value of x, $\quad \Rightarrow \boxed{y = \frac{af-ec}{ad-bc}}$

Now, define a quantity $\boxed{D = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad-bc}$

- note the straight vertical lines, this is not a matrix it is simply a number

Then, $\quad x = \dfrac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$ and $y = \dfrac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$

and $\quad x = \dfrac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{D}, \quad y = \dfrac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{D}$

- D is a 2×2 determinant defined by the coefficient matrix
- To visualise a 2×2 determinant think of 2 arrows and a subtraction i.e.

$$\text{"}\begin{matrix} a & b \\ c & d \end{matrix}\text{"} - \text{"}\begin{matrix} a & b \\ c & d \end{matrix}\text{"} \equiv ad - bc$$

Ex $\quad \begin{matrix} x + 2y = 5 \\ 3x - y = 8 \end{matrix}$ gives $D = \begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix}$

$$= -1 - 6 = -7$$

Then, $\quad x = \dfrac{\begin{vmatrix} 5 & 2 \\ 8 & -1 \end{vmatrix}}{D}, \quad y = \dfrac{\begin{vmatrix} 1 & 5 \\ 3 & 8 \end{vmatrix}}{D}$

i.e. $x = 3, \quad y = 1$

Can you spot the pattern?

$$x = \frac{\text{`D with RHS in column 1'}}{D}, \quad y = \frac{\text{`D with RHS in column 2'}}{D}$$

RHS ≡ right hand side $\longrightarrow$ CRAMER'S RULE $\longleftarrow$

This also works for systems of higher order.

Consider 3 simultaneous equations ....

$$\boxed{\begin{matrix} a_1 x + b_1 y + c_1 z = k_1 \\ a_2 x + b_2 y + c_2 z = k_2 \\ a_3 x + b_3 y + c_3 z = k_2 \end{matrix}}$$

x, y, z are the 3 unknowns and the other symbols denote constants

Then, $\quad x = \dfrac{\begin{vmatrix} k_1 & b_1 & c_1 \\ k_2 & b_2 & c_2 \\ k_3 & b_3 & c_3 \end{vmatrix}}{D}, \quad y = \dfrac{\begin{vmatrix} a_1 & k_1 & c_1 \\ a_2 & k_2 & c_2 \\ a_3 & k_3 & c_3 \end{vmatrix}}{D}$ and

$$z = \frac{\begin{vmatrix} a_1 & b_1 & k_1 \\ a_2 & b_2 & k_2 \\ a_3 & b_3 & k_3 \end{vmatrix}}{D}$$

where now $\quad D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$

but we need to define 3×3 determinants.

In the vector calculus section, I gave the particular case

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$

$$= a_1 (b_2c_3 - c_2b_3) - b_1 (a_2c_3 - c_2a_3) + c_1 (a_2b_3 - b_2a_3)$$

i.e. go along row 1 and, in each case, cover up the row and column of that element to find the 2×2 determinant.

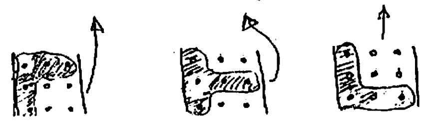
In fact, one can use any row or any column, as long as you keep the signs of each term right. The pattern that gives the signs is $\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$ in the 3×3 case, $\begin{vmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{vmatrix}$ in the 4×4 case,

and so on. Going along the top row, for example, introduces a minus sign for the b_1 term. The signs are $(-1)^{i+j}$.

So, one can also use the first column to get

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}$$

NB.



Terminology:

- The above process is called the "LAPLACE EXPANSION" or the "LAPLACE DEVELOPMENT"
- The 2x2 determinants that result are called "MINORS"
- The "signed minor" is the minor with the appropriate sign and is called the "COFACTOR" (denoted A_{ij})

e.g. For the a_2 term in the last expansion.
a_2 lies in row 2 $(i=2)$ and column 1 $(j=1)$.
The appropriate sign is $(-1)^{i+j} = (-1)^3 = -1$
The minor in this case is $\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix}$ and

the cofactor is

$$A_{21} = -\begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix}$$

The determinant is then $= \sum\limits_{\substack{\text{along row} \\ \text{(or down column)}}} a_{ij}(-1)^{i+j}(\text{minor of } a_{ij})$

$$= \sum_{\text{row/}} a_{ij} A_{ij}$$

This is a general rule for any order of system. **

e.g. for a 4x4 determinant, the Laplace expansion gives minors that are 3x3 determinants. These minors can themselves be expanded by Laplace development to give the result in terms of 2x2 determinants.

— it sounds like a lot of work (and it is!) but it works.

Returning to Cramer's rule...

Ex

$$\begin{aligned} 3x-y-z &= 2 \\ x-2y-3z &= 0 \\ 4x+y+2z &= 4 \end{aligned} \quad \Rightarrow D = \begin{vmatrix} 3 & -1 & -1 \\ 1 & -2 & -3 \\ 4 & 1 & 2 \end{vmatrix} = 2.$$

Then,

$$x = \frac{\begin{vmatrix} 2 & -1 & -1 \\ 0 & -2 & -3 \\ 4 & 1 & 2 \end{vmatrix}}{D} = \frac{2}{2} = 1.$$

$$y = \frac{\begin{vmatrix} 3 & 2 & -1 \\ 1 & 0 & -3 \\ 4 & 4 & 2 \end{vmatrix}}{D} = \frac{4}{2} = 2$$

$$z = \frac{\begin{vmatrix} 3 & -1 & 2 \\ 1 & -2 & 0 \\ 4 & 1 & 4 \end{vmatrix}}{D} = -\frac{2}{2} = -1$$

i.e. solution is $(x,y,z) = (1,2,-1)$.

Here is the details of the workings involved in that last example. To solve

$$\begin{aligned} 3x-y-z &= 2 \\ x-2y-3z &= 0 \\ 4x+y+2z &= 4 \end{aligned},$$

Write the system in matrix form, i.e.

$$\underbrace{\begin{pmatrix} 3 & -1 & -1 \\ 1 & -2 & -3 \\ 4 & 1 & 2 \end{pmatrix}}_{\text{COEFFICIENT MATRIX}} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underbrace{\begin{pmatrix} 2 \\ 0 \\ 4 \end{pmatrix}}_{\text{"RHS"}}.$$

Determinant of the coefficient matrix, D

$$D = \begin{vmatrix} 3 & -1 & -1 \\ 1 & -2 & -3 \\ 4 & 1 & 2 \end{vmatrix} = 3\begin{vmatrix} -2 & -3 \\ 1 & 2 \end{vmatrix} - (-1)\begin{vmatrix} 1 & -3 \\ 4 & 2 \end{vmatrix} + (-1).\begin{vmatrix} 1 & -2 \\ 4 & 1 \end{vmatrix}$$

(expanding along row 1)

i.e. $D = 3.[-4-(-3)] + 1[2-(-12)] - [1-(-8)]$

i.e. $D = 3[-4+3] + [2+12] - [1+8]$

i.e. $D = -3+14-9$

$\therefore D = 2$

"D with RHS in column 1"

$$\begin{vmatrix} 2 & -1 & -1 \\ 0 & -2 & -3 \\ 4 & 1 & 2 \end{vmatrix} = 2\begin{vmatrix} -2 & -3 \\ 1 & 2 \end{vmatrix} - 0 + 4\begin{vmatrix} -1 & -1 \\ -2 & -3 \end{vmatrix}$$

(expanding down column 1 since the zero present simplifies the calculation)

$= 2[-4-(-3)] + 4[3-2]$

$= -2+4 = 2.$

"D with RHS in column 2"

$$\begin{vmatrix} 3 & 2 & -1 \\ 1 & 0 & -3 \\ 4 & 4 & 2 \end{vmatrix} = -2\begin{vmatrix} 1 & -3 \\ 4 & 2 \end{vmatrix} + 0 - 4\begin{vmatrix} 3 & -1 \\ 1 & -3 \end{vmatrix}$$

(expanding down column 2 to exploit the zero, don't forget that the sign table has been used here i.e. $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$)

$= -2[2-(-12)] - 4[-9-(-1)]$

$= -2(14) - 4(-8)$

$= -28+32$

$= 4.$

212

"D with RHS in column 3" $\begin{vmatrix} 3 & -1 & 2 \\ 1 & -2 & 0 \\ 4 & 1 & 4 \end{vmatrix} = 2\begin{vmatrix} 1 & -2 \\ 4 & 1 \end{vmatrix} - 0 + 4\begin{vmatrix} 3 & -1 \\ 1 & -2 \end{vmatrix}$

(using column 3 this time)

$= 2\left[1-(-8)\right] + 4\left[-6-(-1)\right]$

$= 2.(9) + 4.(-5)$

$= 18 - 20$

$= -2$.

Then, $x = \dfrac{\text{"D with RHS in column 1"}}{D} = \dfrac{2}{2} = 1$

$y = \dfrac{\text{"D with RHS in column 2"}}{D} = \dfrac{4}{2} = 2$

$z = \dfrac{\text{"D with RHS in column 3"}}{D} = \dfrac{-2}{2} = -1$

i.e. the solution is $(x,y,z) = (1,2,-1)$.

213

Classification of systems of linear equations

I. Dependence, Consistency, (In)homogeneous-, Singularity

We are now going to introduce some new terminology and classifications regarding systems of simultaneous linear equations.

The purpose of the classifications is to categorise different systems in terms of the character and existence of solutions.

Finally, we will end up with general rules that tell us about the solutions of a system without actually finding the solutions themselves.

These rules are a little abstract. So, in an attempt to provide some insight into their meaning, we will look at types of 2x2 systems where graphical visualisation of the solutions and algebraic manipulation of the equations is relatively straightforward.

214

You will not find this 2x2 development of the terminology and rules in full detail in any books. If you don't like it then fair enough; you can just go straight to the general rules and how to apply them!

Recall how we started this handout by looking at particular 2x2 systems, their graphical interpretation (in terms of the possible intersection of lines), and the nature of their solutions.

Let's try to generalise these ideas ...

In general, 2x2 systems can be written as

$$ax + by = e$$
$$cx + dy = f \quad ,$$

where, unless stated otherwise, we will assume that the constants a,b,c,d and e,f (i.e. those with different symbols) are distinct and non-zero.

- Graphically, these equations can be represented by the lines

$$y = -\frac{a}{b}x + \frac{e}{b}$$

$$y = -\frac{c}{d}x + \frac{f}{d}$$

215

- The equations are called DEPENDENT if one is a multiple of the other. Otherwise, they are called INDEPENDENT.

If our 2x2 system has independent equations then we have $m=2$ independent equations in the $n=2$ unknowns (i.e. x and y).

- If the gradients of the lines are unequal, then there will be a solution (the point where the lines cross). If a solution exists, then the equations are called CONSISTENT. Otherwise, they are called INCONSISTENT.

- One can use the fact of whether a system is HOMOGENEOUS ($e=f=0$) or INHOMOGENEOUS (either $e \neq 0$ or $f \neq 0$, or both e and f non-zero)

216

- The possibilities are that we have :
  - no solution
  - a unique non-trivial solution ($x=y=0$ is the "trivial solution")
  - a unique but trivial solution
  - an infinite number of solutions .

- To classify different cases, one can calculate det(A) , i.e. $|A|$ = the determinant of the coefficient matrix.

If $|A|=0$, then A is said to be <u>SINGULAR</u> .

Note that, if we tried to solve the system using Cramer's rule then , $D=|A|=0$ would give division by zero in the expressions for x and y.

We will now go through the above six topics (marked as "●") for some particular forms of the general 2×2 system and then try to draw conclusions from our findings.

217

The systems that we will consider are:

<u>INHOMOGENEOUS SYSTEMS</u>

<u>Ex 1.</u>
$$\begin{aligned} ax+by &= e \\ cx+by &= f \end{aligned}$$

<u>Ex 2.</u>
$$\begin{aligned} ax+by &= e \\ ax+by &= f \end{aligned}$$

<u>Ex 3.</u>
$$\begin{aligned} ax+by &= e \\ (ma)x+(mb)y &= (me) \end{aligned}$$

[where m is a just a scalar i.e. a number].

<u>HOMOGENEOUS SYSTEMS</u>

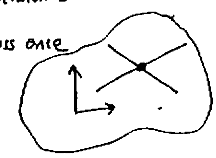
<u>Ex 4.</u>
$$\begin{aligned} ax+by &= 0 \\ cx+by &= 0 \end{aligned}$$

<u>Ex 5.</u>
$$\begin{aligned} ax+by &= 0 \\ (ma)x+(mb)y &= 0 \end{aligned}$$

218

<u>Ex 1.</u> $\boxed{\begin{aligned} ax+by &= e \\ cx+by &= f \end{aligned}} \Rightarrow \begin{aligned} y &= -\frac{a}{b}x+\frac{e}{b} \\ y &= -\frac{c}{b}x+\frac{f}{b} \end{aligned}$

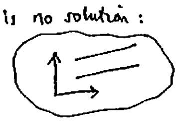
- <u>not</u> a linear multiple (neither $a=c$ nor $e=f$)
  → independent equations , 2 equations in 2 unknowns
- gradients of two lines not equal ⇒ lines cross once
  → unique solution and equations consistent



- equations inhomogeneous ($e\neq0$ and $f\neq0$)
  ∴ $y=x=0$ <u>not</u> a solution
- unique non-trivial solution
- $|A| = ab-bc = b(a-c)\neq0$, since $a\neq c$,
  A non-singular

<u>Ex 2.</u> $\boxed{\begin{aligned} ax+by &= e \\ ax+by &= f \end{aligned}} \Rightarrow \begin{aligned} y &= -\frac{a}{b}x+e/b \\ y &= -\frac{a}{b}+f/b \end{aligned}$

- <u>not</u> a linear multiple since $e\neq f$
  → independent equations , 2 equations in 2 unknowns
- gradients of lines equal but intercepts at $x=0$ different
  → two lines can never cross and there is no solution :
  the equations are inconsistent



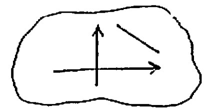
- inhomogeneous system with no solution
- $|A| = ab-ba$ and A is singular (i.e. $|A|=0$)

219

<u>Ex 3.</u> $\boxed{\begin{aligned} ax+by &= e \\ (ma)x+(mb)y &= me \end{aligned}} \Rightarrow \begin{aligned} y &= -\frac{a}{b}x+\frac{e}{b} \\ y &= -\frac{ma}{mb}x\pm\frac{me}{mb} \end{aligned}$

i.e. $\begin{aligned} y &= -\frac{a}{b}x+\frac{e}{b} \\ y &= -\frac{a}{b}x+\frac{e}{b} \end{aligned}$

- One equation is a linear multiple (m) of the other, i.e. the equations are essentially identical and are thus dependent. We only have 1 equation in 2 unknowns
- they are the same equation , any point on this line is a solution , the equations are consistent



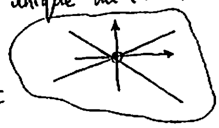
- we have an inhomogeneous system (not supporting the trivial solution ; note that $x=y=0$ <u>requires</u> the RHS constants e and f to be zero) . There are an infinite number of solutions (all the points lying on the line).
- $|A| = \begin{vmatrix} a & b \\ ma & mb \end{vmatrix} = a(mb)-b(ma) = 0$
  i.e. A is singular.

220

Ex. 4 $\boxed{\begin{matrix} ax+by=0 \\ cx+by=0 \end{matrix}}$ Lines are

$$y = -\frac{a}{b}x$$
$$y = -\frac{c}{b}x$$

- Equations independent since $a \neq c$: 2 equations in 2 unknowns (not linear multiple of each other)
- $a \neq c \Rightarrow$ different gradients $\Rightarrow$ intersection at a single solution and the equations are consistent.
- Right hand side is null vector $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$: system homogeneous. Lines thus intersect at the origin, giving unique but trivial solution

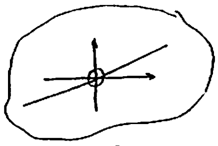


- $|A| = ab - bc = b(a-c) \neq 0$ since $a \neq c$

i.e. $\det A \neq 0$ and A is non-singular

Ex. 5 $\boxed{\begin{matrix} ax+by=0 \\ (ma)x+(mb)y=0 \end{matrix}}$

$$y = -\frac{a}{b}x$$
$$y = -\frac{ma}{mb}x$$



- linear multiple, equations dependent : 1 equation in 2 unknowns
- Any point on line is solution ∴ equations consistent
- Homogeneous with an infinite number of solutions (all points on line)
- $|A| = a(mb) - b(ma) = 0$, A singular

221

Let's tabulate our findings for these $n \times n = 2 \times 2$ systems with m independent equations

INHOMOGENEOUS SYSTEMS (RHS $\neq \underset{\sim}{0}$)

	$ax+by=e$ $cx+by=f$	$ax+by=e$ $ax+by=f$	$ax+by=e$ $(ma)x+(mb)y=me$
dependent/ independent	indept.	indept.	dept.
m indept. equations, m=?	2	2	1
consistent/ inconsistent	consistent	inconsistent	consistent
solutions?	unique non-trivial	no solution	infinite number
det(A)=0 or det(A)≠0	$\lvert A\rvert \neq 0$	$\lvert A\rvert = 0$	$\lvert A\rvert = 0$
A singular or non-singular	non-singular	singular	singular

HOMOGENEOUS SYSTEMS (RHS $= \underset{\sim}{0}$)

	$ax+by=0$ $cx+by=0$	$ax+by=0$ $(ma)x+(mb)y=0$
dependent/ independent	indept.	dependent
m indept. equations, m=?	2	1
consistent/ inconsistent	consistent	consistent
solutions?	unique trivial	infinite number (including trivial)
det(A)=0 or det(A)≠0	$\lvert A\rvert \neq 0$	$\lvert A\rvert = 0$
A singular or non-singular	non-singular	singular

222

Based on these finding, what might be true for $n \times n$ systems (i.e. n equations in n unknowns), where m is the number of independent equations?

<u>$n \times n$ inhomogeneous systems</u>

- If A non-singular ($|A| \neq 0$) and $m = n$ then get A UNIQUE NON-TRIVIAL SOLUTION
- If A singular ($|A| = 0$) and $m < n$ then get AN INFINITE NUMBER OF SOLUTIONS
- If A singular ($|A| = 0$) and $m = n$ then get NO SOLUTION (inconsistency)

---- and for the homogeneous systems ...

223

<u>$n \times n$ homogeneous systems</u>

- Does the trivial solution always exist?

$\longrightarrow$ YES

- A non-singular ($|A| \neq 0$) $\Rightarrow$ ONLY THE TRIVIAL SOLUTION EXISTS.
- A singular ($|A| = 0$) $\Rightarrow$ AN INFINITE NUMBER OF SOLUTIONS (INCLUDING THE TRIVIAL SOLUTION)

HANDOUT 8

CONTENTS

- Classification of systems of linear equations
  II. The Rank of a matrix
- Linear Independence of
  - equations
  - vectors
  - functions

Classification of systems of linear equations

II. The RANK of a matrix

We have now established the terminology, geometry and properties of solutions of 2x2 systems and we have proposed that our findings also apply to all $n \times n$ systems. The catch is that, with higher-order systems, we need a means to determine m (the number of independent equations): this is determined by what is called the RANK of the matrix A. Then, by using information about the right-hand side $\underset{\sim}{b}$, one can work out the existence and number of solutions (as was done for the above 2x2 systems)

Definition

The RANK r of a matrix = the largest non-zero determinant that can be found from its elements (in the order that they appear in the matrix).

In other words, a matrix of rank r contains at least one square sub-matrix of r rows and columns with non-zero determinant while all square sub-matrices of at least $(r+1)$ rows and columns have zero determinant.

Notes

- A square sub-matrix of $\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ could be $\begin{bmatrix} b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}$ or $\begin{bmatrix} a_2 & a_3 \\ b_2 & b_3 \end{bmatrix}$ or even $\begin{bmatrix} a_1 & a_3 \\ b_1 & b_3 \end{bmatrix}$ or $\begin{bmatrix} b_1 & b_3 \\ c_1 & c_3 \end{bmatrix}$. There are a lot to choose from but we would only need to find one with non-zero determinant.
- If A is a square matrix of order n then if the rank of A is n it follows that $|A| \neq 0$ and A is non-singular.

Some Examples

In each case, we will look for the largest subdeterminant that is non-zero.

Ex1 $\begin{bmatrix} 4 & 2 \\ 1 & 5 \end{bmatrix}$ A 2x2 square matrix. The largest square matrix that we can form from this is the matrix itself. So, let's check the determinant of the given 2x2 matrix to see if that is non-zero.

i.e. $\begin{vmatrix} 4 & 2 \\ 1 & 5 \end{vmatrix} = 4.5 - 2.1 = 20 - 2 = 18 \neq 0$

i.e. a matrix with 2 rows and columns with non-zero determinant $\Rightarrow$ matrix has rank $r = 2$.

(227)

Ex2 $\begin{bmatrix}6&3\\8&4\end{bmatrix}$ Again, another square matrix. So check the determinant of this 2x2 matrix.

$$\begin{vmatrix}6&3\\8&4\end{vmatrix} = 6.4-3.8 = 24-24=0. \quad \therefore r<2.$$

Let's now look for square submatrices. There are 1x1 matrices i.e. $[6],[3],[8],[4]$.

A consistent definition of the determinant (in terms of the solution of equations) of a 1x1 matrix is the matrix element itself.

We now try to find a determinant of one of those submatrices that is non-zero. In this case, they all have non-zero determinant. $\therefore$ rank, $r=1$.

Ex3 $\begin{bmatrix}0&0\\0&0\end{bmatrix}$. $\begin{vmatrix}0&0\\0&0\end{vmatrix} = 0-0=0 \quad \therefore r<2$

while all 1x1 submatrices have zero element and thus zero determinant, hence $r<1$.

$\therefore$ rank r of $\begin{bmatrix}0&0\\0&0\end{bmatrix}$ is 0.

(228)

Ex4 $\begin{bmatrix}1&2&8\\4&7&6\\9&5&3\end{bmatrix}$ has $\begin{vmatrix}1&2&8\\4&7&6\\9&5&3\end{vmatrix} = 1.(21-30) - 2(12-54) + 8(20-63)$

$= 9+84-344$

$= -269 \neq 0$

$\therefore$ Matrix has rank $r=3$.

Ex5 $\begin{bmatrix}3&4&5\\1&2&3\\4&5&6\end{bmatrix}$ has $\begin{vmatrix}3&4&5\\1&2&3\\4&5&6\end{vmatrix} = 3(12-15) - 4(6-12) + 5(5-8)$

$= -9+24-15=0$

$\therefore$ rank $r<3$

Now look for a 2x2 submatrix with non-zero determinant.

e.g. test any of $\begin{vmatrix}3&4\\1&2\end{vmatrix}, \begin{vmatrix}4&5\\2&3\end{vmatrix}, \begin{vmatrix}1&2\\4&5\end{vmatrix}, \begin{vmatrix}2&3\\5&6\end{vmatrix}, \ldots$

...we could also test $\begin{vmatrix}3&5\\1&3\end{vmatrix}, \begin{vmatrix}3&5\\4&6\end{vmatrix}, \begin{vmatrix}1&3\\4&6\end{vmatrix}, \begin{vmatrix}3&5\\4&6\end{vmatrix}$, ...etc!

(229)

But we only need to find <u>ONE</u> submatrix with non-zero determinant.

For example, $\begin{vmatrix}3&4\\1&2\end{vmatrix} = 3.2-4.1 = 6-4=2\neq 0$

$\therefore$ rank $r=2$.

We now know how to determine the rank of a single matrix.

Let's try and use this concept to characterise the solution(s) of simultaneous linear equations...

For example, adopting the notation of a_{ij} for the ij^{th} element of coefficient matrix A and b_i for the elements of the right hand side vector, a 3x3 system is written as

$$\boxed{\begin{aligned}a_{11}x+a_{12}y+a_{13}z&=b_1\\a_{21}x+a_{22}y+a_{23}z&=b_2\\a_{31}x+a_{32}y+a_{33}z&=b_3\end{aligned}}$$

i.e. $$\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$$

or simply $A\underset{\sim}{x}=\underset{\sim}{b}$

(230)

One now defines an AUGMENTED COEFFICIENT MATRIX by including $\underset{\sim}{b}$ as a fourth column; let's denote this new matrix by A_b —

i.e. $$A_b = \begin{bmatrix}a_{11}&a_{12}&a_{13}&b_1\\a_{21}&a_{22}&a_{23}&b_2\\a_{31}&a_{32}&a_{33}&b_3\end{bmatrix}.$$

We will now explore the relationship between the ranks of A and A_b and the solution(s) of the equation system.

This is most easily done by considering 2x2 systems.

For a 2x2 system $\boxed{\begin{aligned}ax+by&=e\\cx+dy&=f\end{aligned}}$

one has $A=\begin{bmatrix}a&b\\c&d\end{bmatrix}$ and $A_b=\begin{bmatrix}a&b&e\\c&d&f\end{bmatrix}$.

To see if A_b is of rank 2, one only needs to test $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$, $\begin{vmatrix} b & e \\ d & f \end{vmatrix}$ and $\begin{vmatrix} a & e \\ c & f \end{vmatrix}$.

If these are all zero then A_b will be of rank 1 if it contains a single non-zero element, otherwise it will be of rank 0.

We will now consider the five 2x2 systems listed on p217 and calculate the rank of A and the rank of A_b. The results will be tabulated to see if a pattern emerges regarding the solutions of each case.

(i) $\begin{cases} ax+by=e \\ cx+by=f \end{cases} \rightarrow A=\begin{bmatrix} a & b \\ c & b \end{bmatrix}$ (Recall that we denote elements with different symbols as distinct and non-zero.)

$A_b=\begin{bmatrix} a & b & e \\ c & b & f \end{bmatrix}$

$|A| = ab-bc = b(a-c) \neq 0 \quad \therefore A$ is of rank 2.

Since A_b contains A, A_b is also of rank 2

i.e. $\boxed{\text{rank}(A) = \text{rank}(A_b) = n}$

(where $n=2$ since we are dealing with nxn ie. 2x2 systems)

(ii) $\begin{cases} ax+by=e \\ ax+by=f \end{cases} \rightarrow A=\begin{bmatrix} a & b \\ a & b \end{bmatrix}$

$A_b=\begin{bmatrix} a & b & e \\ a & b & f \end{bmatrix}$

$|A| = ab-ba = 0 \quad \therefore A$ is of rank 1 (elements assumed not all zero)

while $\begin{vmatrix} a & e \\ a & f \end{vmatrix} = af-ea = a(f-e) \neq 0$. $\therefore$ Ab is of rank 2.

i.e. $\boxed{\text{rank}(A) < \text{rank}(A_b) = n}$

(iii) $\begin{cases} ax+by=e \\ (ma)x+(mb)y=me \end{cases} \rightarrow A=\begin{bmatrix} a & b \\ ma & mb \end{bmatrix}$

$A_b=\begin{bmatrix} a & b & e \\ ma & mb & me \end{bmatrix}$

$|A| = a(mb)-b(ma) = 0 \quad \therefore A$ is of rank 1.

While $\begin{vmatrix} a & e \\ ma & me \end{vmatrix} = a(me)-e(ma)=0$

and $\begin{vmatrix} b & e \\ mb & me \end{vmatrix} = b(me)-e(mb)=0$

Since $|A|=0$ also, A_b is also of rank 1.

i.e. $\boxed{\text{rank}(A) = \text{rank}(A_b) < n}$

(iv) $\begin{cases} ax+by=0 \\ cx+by=0 \end{cases} \rightarrow A=\begin{bmatrix} a & b \\ c & b \end{bmatrix}$

$A_b=\begin{bmatrix} a & b & 0 \\ c & b & 0 \end{bmatrix}$

$|A| = ab-bc = b(a-c) \neq 0 \quad \therefore$ both A and A_b rank 2.

ie. $\boxed{\text{rank}(A) = \text{rank}(A_b) = n}$

And finally ...

(v) $\begin{cases} ax+by=0 \\ (ma)x+(mb)y=0 \end{cases} \rightarrow A=\begin{bmatrix} a & b \\ ma & mb \end{bmatrix}$

$A_b=\begin{bmatrix} a & b & 0 \\ ma & mb & 0 \end{bmatrix}$

$|A| = a(mb)-b(ma)=0$ ie. A is of rank 1.

while $\begin{vmatrix} a & 0 \\ ma & 0 \end{vmatrix} = \begin{vmatrix} b & 0 \\ mb & 0 \end{vmatrix} = 0$

$\therefore \boxed{\text{rank}(A) = \text{rank}(A_b) < n}$

Summarise

(i) $\text{rank}(A)=\text{rank}(A_b)=n$	: unique solution
(ii) $\text{rank}(A)<\text{rank}(A_b)=n$	: no solution
(iii) $\text{rank}(A)=\text{rank}(A_b)<n$	: infinite number of sol^{n}s
(iv) $\text{rank}(A)=\text{rank}(A_b)=n$	: unique solution
(v) $\text{rank}(A)=\text{rank}(A_b)<n$	: infinite number of sol^{n}s

→ ONLY 3 DISTINCT CASES

For an nxn system one finds that ...

- $\text{rank}(A)=\text{rank}(A_b)=n \Rightarrow$ UNIQUE SOLUTION
- $\text{rank}(A)=\text{rank}(A_b)<n \Rightarrow$ INFINITE NUMBER OF SOLUTIONS
- $\text{rank}(A)<\text{rank}(A_b) \Rightarrow$ NO SOLUTION

We also know that ...

- Homogeneous systems are always consistent because they always have at least the trivial solution.
- $\det(A)=0$ is required for non-trivial solutions of homogeneous systems.
- Inhomogeneous systems only have non-trivial solutions.

The above six results should be memorised; they tell us the character of the solutions without having to work them out!

What does $\det(A) \equiv |A|$ tell us?

Homogeneous systems : $|A| \neq 0 \Rightarrow$ only trivial solution

$|A| = 0 \Rightarrow$ infinite number of solutions

Inhomogeneous systems : $|A| \neq 0 \Rightarrow$ unique nontrivial solution

$|A| = 0 \Rightarrow$ infinite number of solutions when $r(A) = r(A_b) = m < n$

OR

no solution when $r(A) < r(A_b)$

Note that in Cramer's rule we have

$$x = \frac{|\equiv|}{|A|}, \quad y = \frac{|\equiv|}{|A|}, \text{ etc.}$$

and we cannot find a solution when $|A| = 0$ (using Cramer's rule).

– we cannot divide by zero (when A is singular).

What we have covered so far ...

- consistency and number of solutions of 2×2 systems
  - geometry and algebraic manipulation
- definition of a matrix
- matrix arithmetic
  - addition and subtraction
  - multiplication by a scalar
  - multiplying one matrix by another ("row dot column" rule)
- solutions of equations
  - Cramer's rule
  - Laplace expansion of determinants
  - use of the rank of a matrix

In the very last category, for $n\times n$ systems, denote rank of A by $r(A)$

- $r(A) = r(A_b) = n$ [or simply $r(A)=n$] $\Rightarrow$ unique solution
  m independent equations $(=n)$
  $|A| \neq 0$ (nonsingular A)
- $r(A) = r(A_b) < n \Rightarrow$ infinite number of solutions
  m independent equations $(<n)$
  $|A| = 0$ (singular A)

PLUS For inhomogeneous systems, we can also have inconsistency

i.e. when $r(A) < r(A_b) \Rightarrow$ no solution

$|A| = 0$ (singular A)

This is what we will cover now ...

- clarify what we mean by "linear independence"
  - this connects with rank, singularity, vector spaces and matrices in "echelon form"
- properties of determinants
  - these allow us to simplify determinant calculations
- special types of matrices
  - a brief account of names given to matrices with particular properties
- finding the inverse of a matrix
  - the formal method
  - the row reduction method
- applications of matrices
  - some general contexts where one encounters matrices
- eigenvalues and eigenvectors
  - a very important aspect of matrices

Linear Independence of → equations, → vectors, → functions

Equations

Once again consider the simple 2×2 system

$$x + y = 2 \quad \text{(i)}$$
$$x - y = 0 \quad \text{(ii)}$$

To solve this system by elimination, one may choose to firstly eliminate x from (ii) by subtracting equation (i) from (ii)

i.e. (ii) $\to$ (ii) $-$ (i)

giving $(x - x) - y - y = 0 - 2$

i.e. $-2y = -2$

i.e. $y = 1$

We then have

$$x + y = 2$$
$$y = 1$$

and we can 'back-substitute' to find x

i.e. $x + 1 = 2$, $\therefore x = 1$.

In terms of matrices, we have

$$\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}2\\0\end{pmatrix}$$

i.e. $A\underset{\sim}{x} = \underset{\sim}{b}$

If we manipulate the left hand side of equations, then we have to do the same to the right hand side.

So, let's consider the augmented coefficient matrix

$$A_b = \begin{pmatrix}1&1&2\\1&-1&0\end{pmatrix}\begin{array}{l}\leftarrow \text{row 1} \equiv r_1\\ \leftarrow \text{row 2} \equiv r_2\end{array}$$

To eliminate x from row 2, let $r_2 \to r_2 - r_1$

i.e. $$\begin{pmatrix}1&1&2\\1&-1&0\end{pmatrix} \to \begin{pmatrix}1&1&2\\0&1&1\end{pmatrix} \text{ using } r_2 \to r_2 - r_1$$ "echelon form"

Now consider solving

$$\begin{aligned} x+y &= 2 \quad &(i)\\ 2x+2y &= 4 . \quad &(ii)\end{aligned}$$

$(ii) \to (ii) - 2(i)$ gives

$$\begin{aligned} x+y &= 2 \quad &(i)\\ 0+0 &= 0 \quad &(ii)\end{aligned}$$

i.e. only one equation because (ii) is a linear multiple of (i); in other words, the two equations are not independent — they are dependent.

In terms of matrices, we have

$$\begin{pmatrix}1&1\\2&2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}2\\4\end{pmatrix}$$

and $$A_b = \begin{pmatrix}1&1&2\\2&2&4\end{pmatrix}\begin{array}{l}\leftarrow r_1\\ \leftarrow r_2\end{array}$$

Then, $r_2 \to r_2 - 2r_1$ gives $$\begin{pmatrix}1&1&2\\0&0&0\end{pmatrix}$$ "echelon form"

i.e. a row of zeroes in A_b after trying to eliminate x.

Now consider solving

$$\begin{aligned} x+y &= 2 \quad &(i)\\ x+y &= 5 \quad &(ii)\end{aligned}$$

$(ii) \to (ii) - (i)$ gives

$$\begin{aligned} x+y &= 2\\ 0 &= 3\end{aligned}$$

i.e. an inconsistency

i.e. the equations are inconsistent

i.e. the equations cannot be solved because they do not have a solution.

In terms of matrices, we have

$$\begin{pmatrix}1&1\\1&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}2\\5\end{pmatrix}$$

and $$A_b = \begin{pmatrix}1&1&2\\1&1&5\end{pmatrix}\begin{array}{l}\leftarrow r_1\\ \leftarrow r_2\end{array}$$

Then, $r_2 \to r_2 - r_1$ gives $$\begin{pmatrix}1&1&2\\0&0&3\end{pmatrix}$$ "echelon form"

Now consider a more involved example...

$$\begin{aligned} x + y - 2z &= 7 \quad &(i)\\ 2x - y - 4z &= 2 \quad &(ii)\\ -5x + 4y + 10z &= 1 \quad &(iii)\\ 3x - y - 6z &= 5 \quad &(iv)\end{aligned}$$

Write this system in matrix form (augmented) and eliminate x from equations (ii) to (iv)

i.e. $$A_b = \begin{pmatrix}1&1&-2&7\\2&-1&-4&2\\-5&4&10&1\\3&-1&-6&5\end{pmatrix}\begin{array}{l}\leftarrow r_1\\ \leftarrow r_2\\ \leftarrow r_3\\ \leftarrow r_4\end{array}$$

This can be reduced to $$\begin{pmatrix}1&1&-2&7\\0&-3&0&-12\\0&9&0&36\\0&-4&0&-16\end{pmatrix},$$

using
$$\begin{aligned} r_2 &\to r_2 - 2r_1\\ r_3 &\to r_3 + 5r_1\\ r_4 &\to r_4 - 3r_1\end{aligned}$$

Now eliminate y from the last two equations.

Note that we cannot use multiples of row 1 (ie equation (i)) now, since that would re-introduce x into the last two rows. It is much simpler (and systematic) to use row 2 to eliminate y from the last two equations...

243

i.e. $\begin{pmatrix} 1 & 1 & -2 & 7 \\ 0 & -3 & 0 & -12 \\ 0 & 9 & 0 & 36 \\ 0 & -4 & 0 & -16 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & -2 & 7 \\ 0 & -3 & 0 & -12 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ -- "echelon form"

using $r_3 \to r_3 + 3r_2$

$r_4 \to r_4 - \frac{4}{3} r_2$

We have now reduced the original set of 4 equations to just 2 equations, namely

$$x + y - 2z = 7$$

and $$-3y = -12$$

Note: Had there been an inconsistency in the equations, then a row would have resulted in A_b of the form $\begin{pmatrix} - & - & - & - \\ 0 & 0 & 0 & b \\ - & - & - & - \end{pmatrix}$,

where $b \neq 0$.

244

Let's state formally what we have been doing....

The rules for solving sets of equations by elimination are called the ELEMENTARY ROW OPERATIONS.

i.e. we can
(i) interchange two rows
(ii) multiply a row by a number
(iii) add a multiple of one row to another row

The systematic way to apply elementary row operations is to reduce the matrix to "echelon form". Echelon means "staircase" here and we want to introduce zeroes in the matrix to give the following staircase pattern

$$\begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \\ 0 & 0 & 0 \end{pmatrix}$$ where $*$ are non-zero elements.

In terms of equations, we would then have something like

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

"echelon form"

245

and we can solve this system by working 'upwards' i.e. 'back-substitution',

i.e. $z = b_3/a_{33}$ (we now know z)

then $a_{22} y + a_{23} z = b_2$ (and we can solve to find y)

then $a_{11} x + a_{12} y + a_{13} z = b_1$ (knowing y and z, we can solve this to find x).

In general, for an $n \times n$ system we have...

For consistent equations $\quad \text{rank}(A) = \text{rank}(A_b)$

where the number of non-zero rows of A in echelon form equals m, i.e. number of linearly independent equations

and if $\text{rank}(A) = \text{rank}(A_b) = n$

then we have $m = n$ linearly independent equations and a unique solution

but if $\text{rank}(A) = \text{rank}(A_b) < n$

the we have $m < n$ linearly independent equations and an infinite number of solutions

246

For inconsistent equations $\quad \text{rank}(A) < \text{rank}(A_b)$

i.e. the number of non-zero rows of A in echelon form is less than the number of non-zero rows of A_b in echelon form.

Linear Independence of → equations, vectors, functions

Vectors

- $\underset{\sim}{v_1} = \underset{\sim}{i} + \underset{\sim}{j}$, $\underset{\sim}{v_2} = \underset{\sim}{i} + \underset{\sim}{k}$ and $\underset{\sim}{v_3} = 2\underset{\sim}{i} + \underset{\sim}{j} + \underset{\sim}{k}$ are called linearly dependent because $\underset{\sim}{v_1} + \underset{\sim}{v_2} - \underset{\sim}{v_3} = \underset{\sim}{0}$. We can write each as a linear combination of the other two eg. $\underset{\sim}{v_3} = \underset{\sim}{v_1} + \underset{\sim}{v_2}$.
- $\underset{\sim}{i}$ and $\underset{\sim}{j}$ are linearly independent because $a\underset{\sim}{i} + b\underset{\sim}{j} \neq \underset{\sim}{0}$ always (where a and b are numbers not both zero). i.e. $a\underset{\sim}{i} + b\underset{\sim}{j} = a\begin{pmatrix}1\\0\\0\end{pmatrix} + b\begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}a\\b\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$ only when both a and b are zero.

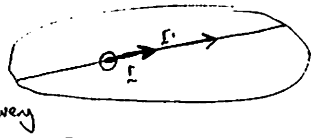
- In general, $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \dots, \underset{\sim}{v}_n$ are linearly dependent if

$a_1\underset{\sim}{v}_1 + a_2\underset{\sim}{v}_2 + \dots + a_n\underset{\sim}{v}_n = \underset{\sim}{0}$ for some numbers $a_1, a_2, \dots, a_n$ that are not all zero.

If their linear combination cannot be set to $\underset{\sim}{0}$ without assuming $a_1 = a_2 = \dots = a_n = 0$ then the vectors are linearly independent.

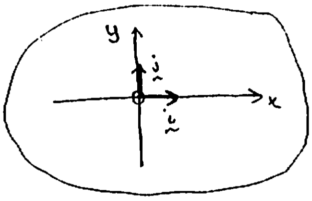
What we are linearly independent vectors?

- 1D space : along a line



a linear multiple of just one vector, say $\underset{\sim}{r}$, can give every position on that line e.g. $\underset{\sim}{r}_1 = 2\underset{\sim}{r}$

- 2D space : the x-y plane

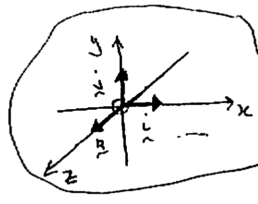


an example of two vectors that are linearly independent is $\underset{\sim}{i} = \begin{pmatrix}1\\0\end{pmatrix}$, $\underset{\sim}{j} = \begin{pmatrix}0\\1\end{pmatrix}$.

whereby, any other vector in this 2D space can be written as a linear combination of these i.e. $\underset{\sim}{r} = x\underset{\sim}{i} + y\underset{\sim}{j}$

$\Rightarrow$ maximum number of linearly independent vectors $= 2$, here.

- 3D space : x, y and z



example of three linearly independent vectors is

$\underset{\sim}{i} = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $\underset{\sim}{j} = \begin{pmatrix}0\\1\\0\end{pmatrix}$, $\underset{\sim}{k} = \begin{pmatrix}0\\0\\1\end{pmatrix}$

any other vector in this 3D space can be written as a linear combination of these i.e. $\underset{\sim}{r} = x\underset{\sim}{i} + y\underset{\sim}{j} + z\underset{\sim}{k}$

$\Rightarrow$ maximum number of linearly independent vectors $= 3$, here

In other words,

the dimension of the space
= the required number of basis vectors
= the number of linearly independent vectors needed to "span" the space

Problems: given a set of vectors, how do we determine if they are linearly independent? If they are not independent, can we determine a smaller set of vectors that are linearly independent?

Solution: in solving sets of equations we take linear combinations of the equations to simplify the system. If we do this with vectors then we can see if they can add up to zero.

OK. So let's put the vectors in rows and pretend that we are solving a set of equations ...

An example Given vectors $\underset{\sim}{v}_1 = (1, 4, -5)$
$\underset{\sim}{v}_2 = (5, 2, 1)$
$\underset{\sim}{v}_3 = (2, -1, 3)$
$\underset{\sim}{v}_4 = (3, -6, 11)$,

are they linearly independent and, if not, can we find a smaller set that is?

Solution Consider the vectors as rows in a matrix (treat them like equations)

i.e. $A = \begin{pmatrix} 1 & 4 & -5 \\ 5 & 2 & 1 \\ 2 & -1 & 3 \\ 3 & -6 & 11 \end{pmatrix}$

$$A = \begin{pmatrix} 1 & 4 & -5 \\ 5 & 2 & 1 \\ 2 & -1 & 3 \\ 3 & -6 & 11 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 4 & -5 \\ 0 & -18 & 26 \\ 0 & -9 & 13 \\ 0 & -18 & 26 \end{pmatrix}$$

using $r_2 \to r_2 - 5r_1$, $r_3 \to r_3 - 2r_1$, $r_4 \to r_4 - 3r_1$

(e.g. row 2 becomes row 2 minus 5 times row 1)

$$\begin{pmatrix} 1 & 4 & -5 \\ 0 & -9 & 13 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

using $r_2 \to r_2 - r_3$, $r_3 \to r_3 + r_2$, $r_4 \to r_4 - r_2$.

Note Get zeroes in column 1 first, then do column 2, etc.

So ...

- there were only two independent vectors in the original four
  i.e. two could be written as a linear combination of the other (independent) two vectors

- $(1, 4, -5)$ and $(0, -9, 13)$ are two linearly independent vectors.

- Note that two linearly independent rows in A is equivalent to saying that $\text{rank}(A) = 2$.

i.e. number of non-zero rows of A in echelon form = rank of A = number of linearly independent row vectors

Functions

Linear (in)dependence of functions is defined in the same way ... Functions $f_1(x), f_2(x), \dots, f_n(x)$ are linearly dependent if some linear combination of them is zero

i.e. $a_1 f_1(x) + a_2 f_2(x) + \dots + a_n f_n(x) = 0$

for some numbers $a_1, a_2, \dots, a_n$ not all zero.

HANDOUT 9

CONTENTS

Properties of determinants

Suppose we need to work out

$$D = \begin{vmatrix} b+c & a & a \\ b & c+a & b \\ c & c & a+b \end{vmatrix}.$$

Laplace expansion gives

$$D = (b+c)\begin{vmatrix} c+a & b \\ c & a+b \end{vmatrix} - a\begin{vmatrix} b & b \\ c & a+b \end{vmatrix} + a\begin{vmatrix} b & c+a \\ c & c \end{vmatrix}$$

$$= (b+c)\left[(a+b)(c+a) - bc\right] - a\left[b(a+b) - bc\right] + a\left[bc - c(c+a)\right]$$

= a lot of work and it's getting really messy....

If we could show that $\begin{vmatrix} b+c & a & a \\ b & c+a & b \\ c & c & a+b \end{vmatrix} = 2\begin{vmatrix} b & a & 0 \\ 0 & c & b \\ c & 0 & a \end{vmatrix}$

then we could work it out in a couple of lines.

i.e. $D = 2\begin{vmatrix} b & a & 0 \\ 0 & c & b \\ c & 0 & a \end{vmatrix} = 2\left[b.(ca) - a(-bc) + 0\right]$

$= \underline{4abc}$

In order to simplify determinants, we need to find the rules of manipulating their elements to ...

- simplify individual elements
- get as many zeroes in there as possible

<u>Let's play about with them and work out the rules</u>

- <u>Multiplying one row (or column) by a number?</u>

e.g. $\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$ and $\Delta' = \begin{vmatrix} Ka & Kb & Kc \\ d & e & f \\ g & h & i \end{vmatrix}$ (K=constant)

→ Laplace expand using the row (or column) with the K factor

$\Rightarrow \Delta' = Ka\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} - Kb\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} + Kc\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}$

$= K\left\{ a\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} - b\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} + c\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}\right\} = \underline{\underline{K\Delta}}$

• What if a whole row (or column) has only zeroes? (253)

e.g. $\Delta = \begin{vmatrix} 0 & 0 & 0 \\ d & e & f \\ g & h & i \end{vmatrix}$

→ Laplace expand across the row (or column) with the zeroes

$\Rightarrow \Delta = 0\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} - 0\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} + 0\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} = 0$

• If two rows are proportional or identical?

e.g. $\Delta = \begin{vmatrix} ka & kb & kc \\ a & b & c \\ d & e & f \end{vmatrix}$

→ Laplace expand along the other row (or column)

$\Rightarrow \Delta = d\,(kbc - kcb) - e\,(kac - kca) + f\,(kab - kba)$

i.e. $\Delta = 0$

If they are identical then just set $k = 1$ to show that $\Delta = 0$

• If elements of a row (or column) are all the sum of two terms?

e.g. $\Delta = \begin{vmatrix} a_1+a_2 & b_1+b_2 & c_1+c_2 \\ d & e & f \\ g & h & i \end{vmatrix}$

(254)

→ Laplace expand along that row or column

$\Rightarrow \Delta = (a_1+a_2)\begin{vmatrix} e & f \\ h & i \end{vmatrix} - (b_1+b_2)\begin{vmatrix} d & f \\ g & i \end{vmatrix} + (c_1+c_2)\begin{vmatrix} d & e \\ g & h \end{vmatrix}$

$= \left\{a_1\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} - b_1\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} + c_1\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}\right\} + \left\{a_2\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} - b_2\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix} + c_2\begin{vmatrix} \cdot & \cdot \\ \cdot & \cdot \end{vmatrix}\right\}$

i.e. $\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ d & e & f \\ g & h & i \end{vmatrix} + \begin{vmatrix} a_2 & b_2 & c_2 \\ d & e & f \\ g & h & i \end{vmatrix}$

• Adding a multiple of one row (or column) to another row (or column)?

e.g. $\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$ and $\Delta' = \begin{vmatrix} a+kd & b+ke & c+kf \\ d & e & f \\ g & h & i \end{vmatrix}$

Use the previous result (top row has elements as the sum of two terms)

$\Rightarrow \Delta' = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} - \underbrace{\begin{vmatrix} kd & ke & kf \\ d & e & f \\ g & h & i \end{vmatrix}}_{\text{one row is a multiple of another} \Rightarrow \text{equals zero}}$

$\Rightarrow$ determinant remains unchanged.

What if two rows (or two columns) are interchanged? (255)

e.g. $\Delta = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$ and $\Delta' = \begin{vmatrix} d & e & f \\ a & b & c \\ g & h & i \end{vmatrix}$

$= a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$. $\quad \Delta' = -a\begin{vmatrix} e & f \\ h & i \end{vmatrix} + b\begin{vmatrix} d & f \\ g & i \end{vmatrix} - c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$

(expanding along row 1) $\quad \underline{\Delta' = -\Delta} \quad$ (expanding along row 2)

So, it follows from the sign table " $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$ " that interchanging 2 rows (or columns) gives $\Delta \to -\Delta$.

• What if we take the transpose of the matrix?

The transpose is when the rows are taken as columns and the columns are taken as rows : $A \to A^T$ ("A transpose")

e.g. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then $A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$

e.g. $A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ then $A^T = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$, etc.

In the 2x2 case $|A| = ad - bc$, $|A^T| = ad - cb = |A|$.

This property can be used to prove the 3x3 case and then that can be used to prove the 4x4 case, and so on. $\boxed{|A| = |A^T|}$

Summarise ... (256)

Properties of determinants

1) each element of one row (or one column) times k $\Rightarrow \Delta \to k\Delta$

2) $\Delta = 0$ if (a) all elements in one row (or column) zero
or (b) two rows (or columns) identical or proportional

3) Interchanging two rows (or two columns) $\Rightarrow \Delta \to -\Delta$

4) Δ unchanged if (a) $A \to A^T$
or (b) add multiples of one row (or column) to another

— the above four properties need to be memorised.

Ex. $D = \begin{vmatrix} 4 & 3 & 0 & 1 \\ 9 & 7 & 2 & 3 \\ 4 & 0 & 2 & 1 \\ 3 & -1 & 4 & 0 \end{vmatrix}$

$= -2\begin{vmatrix} 1 & 3 & 0 & 4 \\ 3 & 7 & 1 & 9 \\ 1 & 0 & 1 & 4 \\ 0 & -1 & 2 & 3 \end{vmatrix}$ Interchanging columns 1 and 4 ($c_1 \leftrightarrow c_4 \Rightarrow$ minus sign). Factor out 2 from column 3.

$= -2\begin{vmatrix} 1 & 3 & 0 & 4 \\ 0 & -2 & 1 & -3 \\ 0 & -3 & 1 & 0 \\ 0 & -1 & 2 & 3 \end{vmatrix}$ $\quad r_2 \to r_2 - 3r_1$, $\quad r_3 \to r_3 - r_1$

$$= -2\begin{vmatrix} -2 & 1 & -3 \\ -3 & 1 & 0 \\ -1 & 2 & 3 \end{vmatrix}$$ Laplace expansion down column 1 gives only one term that is non-zero

$$= -2\begin{vmatrix} 1 & 2 & -3 \\ 1 & 3 & 0 \\ 2 & 1 & 3 \end{vmatrix}$$ factor out -1 from column 1 $\Rightarrow$ minus sign; $c_1 \leftrightarrow c_2$ (another minus sign): two minus signs cancel out

$$= -2\begin{vmatrix} 1 & 2 & -3 \\ 0 & 1 & 3 \\ 0 & -3 & 9 \end{vmatrix}$$ $r_2 \to r_2 - r_1$, $r_3 \to r_3 - 2r_1$

$$= -2\begin{vmatrix} 1 & 3 \\ -3 & 9 \end{vmatrix} = -2(9+9) = -36.$$ {Where a Laplace expansion down column 1 gives a single non-zero term}

Special types of matrices (I)

- The transpose, A^T : interchange rows and columns in A
- The complex conjugate, A^* : take the complex conjugate of each element
- Symmetric matrix, $A = A^T$ e.g. $A = \begin{pmatrix} a & b & c \\ b & e & d \\ c & d & f \end{pmatrix}$
- Hermitian matrix, $A = A^\dagger$ (read as "A dagger") where $A^\dagger = (A^T)^*$

- The unit matrix I = square matrix full of zeroes except for the main diagonal which has 1's

e.g. $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, etc

The unit matrix plays a role in matrix algebra similar to that played by the number 1 in ordinary algebra

e.g. $IA = AI = A$

or $II = I$, $I^n = I$ $(n = 1, 2, 3, \ldots)$

Finding the inverse of a matrix

The inverse of $n \times n$ matrix A is another $n \times n$ matrix A^{-1} such that

$$\boxed{AA^{-1} = A^{-1}A = I}.$$

- If A does not have an inverse then it is singular
- $II = I \Rightarrow I$ is its own inverse

Ex if $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

then $A^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & -\frac{1}{2} \end{bmatrix}$

since $AA^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ 3/2 & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Ex if $A = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$ (a diagonal matrix)

then $A^{-1} = \begin{bmatrix} 1/\lambda_1 & 0 & \cdots & 0 \\ 0 & 1/\lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1/\lambda_n \end{bmatrix}$

since $AA^{-1} = \begin{bmatrix} \lambda_1/\lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2/\lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n/\lambda_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}$

We have verified a couple of examples by direct multiplication but how do we find A^{-1}?

There are actually a number of ways. I will outline two of these — the formal method

— the row reduction method

The Formal Method

This can be proved by using Cramer's method, but we will just state the result here.

Recall the terminology used in Laplace expansion ...

... in calculating $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$, consider expansion along row 1 and the a_{12} term

i.e the cofactor of a_{ij} is the signed minor of a_{ij}

for a_{12}, the cofactor is $A_{ij} = (-1)^{i+j}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$

i.e. $A_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$

Now consider a matrix filled with the cofactors of each element eg. $C = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$

The inverse, $\boxed{A^{-1} = \frac{1}{|A|}C^T}$ **

C^T is called the adjoint of A.

- Note that for A^{-1} to exist A must be square
  – for example $\det A = |A|$ is only defined for square matrices

- Given A^{-1}, we can solve the system $\boxed{A\underset{\sim}{x} = \underset{\sim}{b}}$

i.e. $A^{-1}A\underset{\sim}{x} = A^{-1}\underset{\sim}{b}$

i.e. $I\underset{\sim}{x} = A^{-1}\underset{\sim}{b}$

i.e. $\boxed{\underset{\sim}{x} = A^{-1}\underset{\sim}{b}}$

Ex
$$\underbrace{\begin{bmatrix} 5 & 8 & 1 \\ 0 & 2 & 1 \\ 4 & 3 & -1 \end{bmatrix}}_{A}\underbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}_{\underset{\sim}{x}} = \underbrace{\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}}_{\underset{\sim}{b}}$$
Calculate A^{-1} and hence solve for $\underset{\sim}{x}$.

Outline Ans The matrix of cofactors for A is found (after calculating all the appropriate 2x2 determinants with their signs given by the sign table)...
$$C = \begin{bmatrix} -5 & +4 & -8 \\ +11 & -9 & +17 \\ -6 & -5 & +10 \end{bmatrix}$$

$|A| = 5(-5) + 4(6) = -1$ (down 1st column)

$$\rightarrow A^{-1} = \frac{1}{|A|}C^T = -1.\begin{bmatrix} -5 & 11 & -6 \\ 4 & -9 & -5 \\ -8 & 17 & 10 \end{bmatrix} = \begin{bmatrix} -5 & -11 & 6 \\ -4 & 9 & 5 \\ 8 & -17 & -10 \end{bmatrix}$$

$$\rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = A^{-1}\underset{\sim}{b} = \begin{bmatrix} -5 & -11 & 6 \\ -4 & 9 & 5 \\ 8 & -17 & -10 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 3 \end{bmatrix}$$
ie. $x = 3$, $y = -2$, $z = 3$

Full Ans

The matrix of cofactors of A, $C = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$

where $A = \begin{bmatrix} 5 & 8 & 1 \\ 0 & 2 & 1 \\ 4 & 3 & -1 \end{bmatrix} \Rightarrow A_{11} = +\begin{vmatrix} 2 & 1 \\ 3 & -1 \end{vmatrix} = -2-3 = -5$

recall sign table $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$

$A_{21} = -\begin{vmatrix} 8 & 1 \\ 3 & -1 \end{vmatrix} = -(-8-3) = +11$

$A_{31} = +\begin{vmatrix} 8 & 1 \\ 2 & 1 \end{vmatrix} = 8-2 = 6$

$A_{12} = -\begin{vmatrix} 0 & 1 \\ 4 & -1 \end{vmatrix} = -(0-4) = +4$, $\quad A_{13} = +\begin{vmatrix} 0 & 2 \\ 4 & 3 \end{vmatrix} = 0-8 = -8$

$A_{22} = +\begin{vmatrix} 5 & 1 \\ 4 & -1 \end{vmatrix} = -5-4 = -9$, $\quad A_{23} = -\begin{vmatrix} 5 & 8 \\ 4 & 3 \end{vmatrix} = -(15-32) = +17$

$A_{32} = -\begin{vmatrix} 5 & 1 \\ 0 & 1 \end{vmatrix} = -(5-0) = -5$, $\quad A_{33} = +\begin{vmatrix} 5 & 8 \\ 0 & 2 \end{vmatrix} = 10-0 = 10$

$$\Rightarrow C = \begin{bmatrix} -5 & 4 & -8 \\ 11 & -9 & 17 \\ 6 & -5 & 10 \end{bmatrix} \text{ and } C^T = \begin{bmatrix} -5 & 11 & 6 \\ 4 & -9 & -5 \\ -8 & 17 & 10 \end{bmatrix}$$
(swapping rows with columns to get the transpose)

$\det(A) = 5.A_{11} + 0.A_{21} + 4.A_{31}$ (going down 1st column to exploit the zero term)

$= 5\begin{vmatrix} 2 & 1 \\ 3 & -1 \end{vmatrix} + 0.\left\{-\begin{vmatrix} 8 & 1 \\ 3 & -1 \end{vmatrix}\right\} + 4\begin{vmatrix} 8 & 1 \\ 2 & 1 \end{vmatrix}$ (i.e. Laplace expansion)

$= 5.(-2-3) + 0 \quad + 4(8-2)$

$= -25+24$. i.e. $\det(A) = |A| = -1$

$$\Rightarrow A^{-1} = \frac{1}{|A|}C^T = -\begin{pmatrix} -5 & 11 & 6 \\ 4 & -9 & -5 \\ -8 & 17 & 10 \end{pmatrix} = \begin{pmatrix} 5 & -11 & -6 \\ -4 & 9 & 5 \\ 8 & -17 & -10 \end{pmatrix}$$

$$\Rightarrow \underset{\sim}{x} = A^{-1}\underset{\sim}{b} = \begin{pmatrix} 5 & -11 & -6 \\ -4 & 9 & 5 \\ 8 & -17 & -10 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 10+11-18 \\ -8-9+15 \\ 16+17-30 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ 3 \end{pmatrix}$$
i.e. $x = 3$, $y = -2$, $z = 3$

The Row Reduction Method

≡ "by applying elementary row operations to the identity matrix"

Consider solving $\boxed{A\underset{\sim}{x} = \underset{\sim}{b}}$ for $\underset{\sim}{x}$.

This system is equivalent to $\boxed{A\underset{\sim}{x} = I\underset{\sim}{b}}$

$$\text{ie. } \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

If A^{-1} exists, we can apply elementary row operations to both the left and right hand side of this equation to reduce it to

$$\underset{\sim}{x} = A^{-1}I\underset{\sim}{b}$$

or, equivalently, $\boxed{I\underset{\sim}{x} = A^{-1}\underset{\sim}{b}}$.

Consider each elementary row operation as premultiplying the system with an appropriate matrix E_k ($k=1,..,m$ for m elementary row operations).

We then have...
$$E_mE_{m-1}\cdots E_2E_1A\underset{\sim}{x} = E_mE_{m-1}\cdots E_2E_1I\underset{\sim}{b}$$
in order to reduce the system to...
$$I\underset{\sim}{x} = A^{-1}\underset{\sim}{b}.$$

i.e. by transforming $A \to I$ we get $\boxed{A^{-1} = E_mE_{m-1}\cdots E_2E_1I}$ **

To simultaneously apply elementary row operations to both the left and right hand sides, it is convenient to form the "combined coefficient matrix": $\boxed{A \mid I}$.

from left hand side — from right hand side

One then applies $E_1, E_2, .., E_m$ to the combined matrix until it becomes $\boxed{I \mid A^{-1}}$ then we have found A^{-1}.

Ex Solve the system
$$\begin{aligned} 2x_1 + x_2 + x_3 &= 5 \\ x_1 + 3x_2 + 2x_3 &= 1 \\ 3x_1 - 2x_2 - 4x_3 &= -4 \end{aligned}$$

Ans Combined coefficient matrix:
$$\left[\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 3 & 2 & 0 & 1 & 0 \\ 3 & -2 & -4 & 0 & 0 & 1 \end{array}\right]$$

The goal is to reduce A to I using elementary row operations

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 2 & 0 & 1 & 0 \\ 2 & 1 & 1 & 1 & 0 & 0 \\ 3 & -2 & -4 & 0 & 0 & 1 \end{array}\right]$$
by interchanging rows 1 and 2 to get the useful "1" in the top left hand corner (i.e. $r_1 \leftrightarrow r_2$)

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & -5 & -3 & 1 & -2 & 0 \\ 0 & -11 & -10 & 0 & -3 & 1 \end{array}\right]$$
where $r_2 \to r_2 - 2r_1$, $r_3 \to r_3 - 3r_1$ (we are now using that "1")

$$\rightarrow \left[\begin{array}{ccc|ccc} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & -5 & -3 & 1 & -2 & 0 \\ 0 & -1 & -4 & -2 & 1 & 1 \end{array}\right]$$
where $r_3 \to r_3 - 2r_2$ (just to simplify row 3)

$\triangleright \left[\begin{array}{ccc|ccc} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & 1 & 4 & 2 & -1 & -1 \\ 0 & 5 & 3 & -1 & 2 & 0 \end{array}\right]$ $r_2 \to -r_2$, $r_3 \to -r_3$ {to reduce amount of negative numbers}
$r_2 \leftrightarrow r_3$ {get the useful "1" in the diagonal}

$\triangleright \left[\begin{array}{ccc|ccc} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & 1 & 4 & 2 & -1 & -1 \\ 0 & 0 & -17 & -11 & 7 & 5 \end{array}\right]$ $r_3 \to r_3 - 5r_2$ {use that "1"}

$\triangleright \left[\begin{array}{ccc|ccc} 1 & 0 & -10 & -6 & 4 & 3 \\ 0 & 1 & 4 & 2 & -1 & -1 \\ 0 & 0 & 1 & 11/17 & -7/17 & -5/17 \end{array}\right]$ $r_1 \to r_1 - 3r_2$ {use that "1" to get zero on top row}
$r_3 \to r_3/(-17)$ {get another "1" in the diagonal}

$\triangleright \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 8/17 & -2/17 & 1/17 \\ 0 & 1 & 0 & -10/17 & 11/17 & 3/17 \\ 0 & 0 & 1 & 11/17 & -7/17 & -5/17 \end{array}\right]$ $r_1 \to r_1 + 10r_3$, $r_2 \to r_2 - 4r_3$ {use the "1" to finish off the column on the left hand side}

we now have the form $I \,\vdots\, A^{-1}$

i.e. $A^{-1} = \frac{1}{17}\begin{bmatrix} 8 & -2 & 1 \\ -10 & 11 & 3 \\ 11 & -7 & -5 \end{bmatrix}$. Then, $\underset{\sim}{x} = A^{-1}\underset{\sim}{b}$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{17}\begin{bmatrix} 8 & -2 & 1 \\ -10 & 11 & 3 \\ 11 & -7 & -5 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \\ -4 \end{bmatrix} = \frac{1}{17}\begin{bmatrix} 34 \\ -51 \\ 68 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}$$

$$\Rightarrow \underline{\underline{x_1 = 2,\ x_2 = -3,\ x_3 = 4}}$$

More special matrices (II)

- An orthogonal matrix has $\boxed{A^{-1} = A^T}$ i.e. $A^T A = I$
- A unitary matrix has $\boxed{A^{-1} = (A^T)^*}$ i.e. $(A^T)^* A = I$

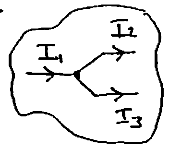
Note: a real unitary matrix is orthogonal.

Applications of Matrices

(1) Solution of simultaneous linear equations

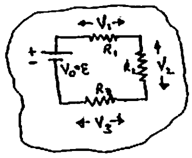
e.g. Kirchoff's laws

- sum of currents into junction = 0



$\Rightarrow I_1 - I_2 - I_3 = 0$

- sum of potential differences around a closed circuit = 0



$\Rightarrow V_0 - V_1 - V_2 - V_3 = 0$

a more complex example, resulting in a larger system of equations, is analysis of a "Wheatstone Bridge"... (don't worry about detail)

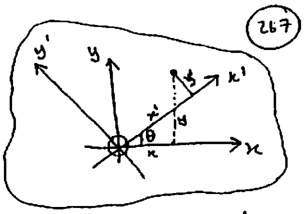
$$\begin{aligned} \rightarrow\ (R_3+R_4)I_1 \quad & -R_3 I_2 \quad & -R_4 I_3 = V \\ R_3 I_1 - (R_1+R_3+R_5)I_2 \quad & & +R_5 I_3 = 0 \\ R_4 I_1 \quad & +R_5 I_2 - (R_2+R_4+R_5)I_3 = 0 & \end{aligned}$$

$$\rightarrow \begin{bmatrix} R_3+R_4 & -R_3 & -R_4 \\ R_3 & -(R_1+R_3+R_5) & R_5 \\ R_4 & R_5 & -(R_2+R_4+R_5) \end{bmatrix}\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} V \\ 0 \\ 0 \end{bmatrix}$$

For example, use Cramer's rule to solve for I_1, I_2, I_3.

- There are many, many more areas of physics where the solution of simultaneous linear equations arises.

2) Rotation of a reference frame



Say, by angle θ whereby coordinates $(x, y) \to (x', y')$

Then, $x' = x\cos\theta + y\sin\theta$
$y' = -x\sin\theta + y\cos\theta$ i.e. $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix}$

A ≡ "ROTATION MATRIX"

i.e. $\boxed{\underset{\sim}{r}' = A\underset{\sim}{r}}$ where $\underset{\sim}{r}' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ and $\underset{\sim}{r} = \begin{pmatrix} x \\ y \end{pmatrix}$

On geometrical grounds, the inverse matrix will correspond to a rotation of $-\theta$ i.e. $\boxed{\underset{\sim}{r} = A^{-1}\underset{\sim}{r}'}$

$$\text{where } A^{-1} = \begin{pmatrix} \cos(-\theta) & \sin(-\theta) \\ -\sin(-\theta) & \cos(-\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

You can check this by multiplication i.e. that $AA^{-1} = I$.

A further rotation ϕ of axes (x', y') to (x'', y'') is then given by

$\boxed{\underset{\sim}{r}'' = B\underset{\sim}{r}' = BA\underset{\sim}{r}}$ where $B = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$ and $\underset{\sim}{r}'' = \begin{pmatrix} x'' \\ y'' \end{pmatrix}$.

- Note. The combined operation of "A then B" results in a matrix product BA equivalent to a rotation $\theta + \phi$.

(3) General linear transformations

One can define more general linear transformations (instead of just rotations) to transform a vector $\underset{\sim}{r}$ in the xyz-system to another $\underset{\sim}{r}'$ in the $x'y'z'$-system...

i.e. $\boxed{\underset{\sim}{r}' = A\underset{\sim}{r}}$ where $A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$

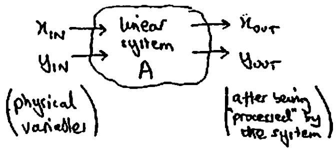
and $\underset{\sim}{r}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$, $\underset{\sim}{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

i.e. $x' = a_1x + a_2y + a_3z$
$y' = b_1x + b_2y + b_3z$
$z' = c_1x + c_2y + c_3z$

- One can also combine such transformations as matrix products, say BA, and use the combined transformation matrix rather than operating with A and then with B.
- Of course, to transform from $\underset{\sim}{r}'$ back to $\underset{\sim}{r}$, say, then use the inverse transformation which is equivalent to A^{-1}
- To transform from $\underset{\sim}{r}''$ back to $\underset{\sim}{r}$, where $\underset{\sim}{r}'' = BA\underset{\sim}{r}$, then $B^{-1}\underset{\sim}{r}'' = A\underset{\sim}{r}$ and $A^{-1}B^{-1}\underset{\sim}{r}'' = \underset{\sim}{r}$

i.e. $\boxed{(BA)^{-1} = A^{-1}B^{-1}}$ A general result.

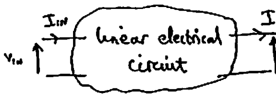
I) Systems Theory



Linear system : $x_{OUT} = a\,x_{IN} + b\,y_{IN}$
$y_{OUT} = c\,x_{IN} + d\,y_{IN}$

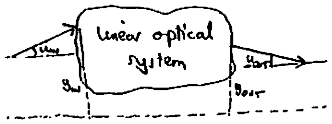
i.e. $\begin{pmatrix} x_{OUT} \\ y_{OUT} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_{IN} \\ y_{IN} \end{pmatrix}$

e.g.



(voltages and currents)

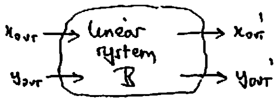
e.g.



(tracing a light ray in terms of its height y and direction u)

We can then send the output through another linear system $B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$ and get x_{OUT}', y_{OUT}'

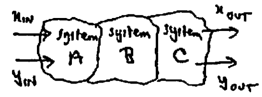
i.e.



where $\begin{pmatrix} x_{OUT}' \\ y_{OUT}' \end{pmatrix} = B \begin{pmatrix} x_{OUT} \\ y_{OUT} \end{pmatrix} = BA \begin{pmatrix} x_{IN} \\ y_{IN} \end{pmatrix}$

- systems are "cascaded" to give overall transmission of $\begin{pmatrix} x_{IN} \\ y_{IN} \end{pmatrix} \rightarrow \begin{pmatrix} x_{OUT}' \\ y_{OUT}' \end{pmatrix}$

And so on...

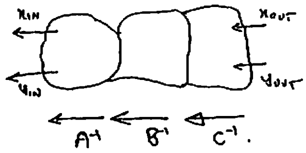


$\begin{pmatrix} x_{OUT} \\ y_{OUT} \end{pmatrix} = CBA \begin{pmatrix} x_{IN} \\ y_{IN} \end{pmatrix}$

$= D \begin{pmatrix} x_{IN} \\ y_{IN} \end{pmatrix}$, say.

Then, $\boxed{D^{-1} = A^{-1}B^{-1}C^{-1}}$

i.e.



and $\begin{pmatrix} x_{IN} \\ y_{IN} \end{pmatrix} = A^{-1}B^{-1}C^{-1} \begin{pmatrix} x_{OUT} \\ y_{OUT} \end{pmatrix}$

In the above systems examples, the matrices are used to describe a discrete "jump" (in time or space or both) through the system.

i.e. the output is given directly in terms of the input and we do not know how much time it took to go through the system or the details of the "path" in space that was taken.

- Matrices can also be used to describe continuous spatial or temporal evolution of system parameters (x and y, say). Consider, for example, temporal evolution (evolution in time t) governed by a pair of ordinary differential equations.

e.g. $$\boxed{\begin{aligned} \frac{dx}{dt} &= 2x + 3y \\ \frac{dy}{dt} &= 2x + y \end{aligned}}$$

i.e. $\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+3y \\ 2x+y \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

ie. $\boxed{\frac{d\underset{\sim}{u}}{dt} = M\underset{\sim}{u}}$ where $\underset{\sim}{u} = \begin{pmatrix} x \\ y \end{pmatrix}$, $M = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$.

This is a very compact way of writing what could be a very large system of equations.

e.g. $\underset{\sim}{u}$ could be an n-dimensional vector where n is large and M an $n \times n$ matrix.

Eigenvalues and Eigenvectors

In the above examples of rotation, general linear transformation, linear systems and differential equations, a matrix is used to describe the discrete or continuous change of a vector.

A very important class of problem that arises in many areas of physics is determining the vectors $\underset{\sim}{x}$ that do not change direction under transformation by matrix A (the eigenvectors of A).

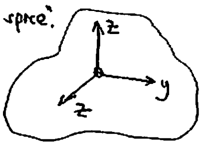
In this case, we have $\boxed{A\underset{\sim}{x} = \lambda \underset{\sim}{x}}$

where λ is called the eigenvalue of A corresponding to the eigenvector $\underset{\sim}{x}$. The eigenvalues and eigenvectors of A are also called the characteristic values and vectors of A, respectively.

In physical problems, they can define the characteristic or "normal" modes of a system e.g. the normal modes of vibration. In quantum mechanics, they are often called the "eigenstates".

More complex evolution in the system (when it's not in an eigenstate) can be written in terms of a complete set of eigenvectors. The eigenvectors of a system are linearly independent.

If a solution is three-dimensional, $\underset{\sim}{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ for example, then all possible solutions of the "system" A (a 3×3 matrix which could be equivalent to a differential equation e.g. $\frac{d}{dt}\underset{\sim}{x} = A\underset{\sim}{x}$) lie somewhere in the three-dimensional "solution space".



If we have the three linearly independent eigenvectors $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \underset{\sim}{x}_3$ of A, then any solution in this three-dimensional space can be written as a linear combination of the eigenvectors. They act like "basis vectors" for the system.

→ How do we find the eigenvalues and eigenvectors of

$$A\underset{\sim}{x} = \lambda \underset{\sim}{x} \quad ?$$

For an $n \times n$ system, we would have

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{bmatrix}.$$

$$\leftarrow A \rightarrow \quad \underset{\sim}{x} = \lambda \underset{\sim}{x}$$

Recall that this actually means ...

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= \lambda x_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= \lambda x_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= \lambda x_n \end{aligned}$$

i.e.
$$\begin{aligned} (a_{11}-\lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + (a_{22}-\lambda)x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn}-\lambda)x_n &= 0 \end{aligned}$$

i.e.
$$\begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Another way to write this would be in terms of the identity matrix I:

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad \text{then} \quad \lambda I = \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix}$$

so
$$A'\underset{\sim}{x} = \underset{\sim}{0}$$
where $A' = A - \lambda I$

$A'\underset{\sim}{x} = \underset{\sim}{0}$ is just a homogeneous system and we thus have only two cases ...

(i) $\det A' \neq 0$ ⇒ only the trivial solution $\underset{\sim}{x} = \underset{\sim}{0}$

... not very useful. It doesn't tell us anything about the original system A; all homogeneous systems have this solution.

(ii) $\det A' = 0$ ⇒ an infinite number of solutions

... it is the form of these solutions that give us the eigenvectors (the characteristic modes).

In other words, we seek eigenvalues λ that give us a singular homogeneous system i.e. we solve

$$\boxed{\det(A - \lambda I) = 0}$$ THE CHARACTERISTIC EQUATION

to find the eigenvalues.

Ex A is a 2×2 matrix $\det(A-\lambda I)=0$ gives

$$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = 0 \quad \text{i.e.} \quad (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21} = 0$$

This is of the form $\boxed{\lambda^2 + c_1\lambda + c_0 = 0}$ (c_0, c_1 constants).

⇒ two possible eigenvalues λ_1 and λ_2.

Ex A is a 3×3 matrix $\det(A-\lambda I)=0$ gives

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix} = 0$$

i.e.
$$\underbrace{(a_{11}-\lambda)\underbrace{\begin{vmatrix} a_{22}-\lambda & a_{23} \\ a_{32} & a_{33}-\lambda \end{vmatrix}}_{\text{quadratic in } \lambda}}_{\text{cubic in } \lambda} - a_{12}\underbrace{\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33}-\lambda \end{vmatrix}}_{\text{linear in } \lambda} + a_{13}\underbrace{\begin{vmatrix} a_{21} & a_{22}-\lambda \\ a_{31} & a_{32} \end{vmatrix}}_{\text{linear in } \lambda} = 0$$

This is of the form $\boxed{\lambda^3 + c_2\lambda^2 + c_1\lambda + c_0 = 0}$

⇒ three possible eigenvalues λ_1, λ_2 and λ_3.

Ex A is an $n \times n$ matrix $\det(A-\lambda I)=0$ gives

a characteristic equation of the form: $\boxed{\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0 = 0}$

⇒ n possible eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Given an eigenvalue λ_1, how do we find the corresponding eigenvector $\underset{\sim}{x}_1$?

→ Substitute for λ_1 and solve $A\underset{\sim}{x}_1 = \lambda_1 \underset{\sim}{x}_1$.

Ex. $A = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$. Then $|A - \lambda I| = 0$ yields $\begin{vmatrix} 2-\lambda & 3 \\ 4 & 1-\lambda \end{vmatrix} = 0$

$\Rightarrow (2-\lambda)(1-\lambda) - 12 = 0$

$\Rightarrow \lambda^2 - 3\lambda - 10 = 0$

$\Rightarrow \lambda_1 = -2$ and $\lambda_2 = 5$

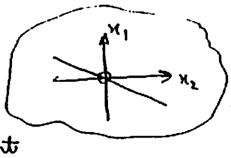
Now, $(A - \lambda I)\underset{\sim}{x} = \underset{\sim}{0}$ gives $\left\{ \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right\} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

i.e. $\begin{pmatrix} 2-\lambda & 3 \\ 4 & 1-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\lambda_1 = -2$ gives $\begin{pmatrix} 4 & 3 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Recall that we forced the coefficient matrix to be singular.

i.e. $4x_1 + 3x_2 = 0$



This is just a line in the x_1-x_2 plane. We cannot specify x_1 and x_2 but we know that they are characterised by the gradient

i.e. $x_1/x_2 = -3/4$. The eigenvalue is then $\boxed{\underset{\sim}{x}_1 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \alpha \begin{pmatrix} 3 \\ -4 \end{pmatrix}}$

where α is an undetermined scalar (number)

Note that if $A\underset{\sim}{x}_1 = \lambda_1 \underset{\sim}{x}_1$

then $A\alpha\underset{\sim}{x}_1 = \lambda_1 \alpha \underset{\sim}{x}_1$

i.e. if, for λ_1, $\underset{\sim}{x}_1$ is an eigenvector then so is $\alpha \underset{\sim}{x}_1$.

$\lambda_2 = 5$ gives $\begin{pmatrix} 2-\lambda_2 & 3 \\ 4 & 1-\lambda_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

i.e. $\begin{pmatrix} -3 & 3 \\ 4 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\Rightarrow -3x_1 + 3x_2 = 0$ and $x_1/x_2 = 1$.

The second eigenvector is then $\underset{\sim}{x}_2 = \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

where β is another undetermined scalar.

Another example

Consider the time evolution of variables x and y that is governed by the differential equations

$$\boxed{\begin{aligned} \frac{dx}{dt} &= 2x + 3y \\ \frac{dy}{dt} &= 2x + y \end{aligned}}$$

Write $\underset{\sim}{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ and $A = \begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix}$ to give $\frac{d}{dt}\underset{\sim}{x} = A\underset{\sim}{x}$

* To find the eigenvalues of the system (corresponding to the eigenvectors of the system that preserve their direction as time evolves .... note that since eigenvectors are only defined in terms of their direction, when a system is in an eigenstate then it remains in that eigenstate) .... *

set $|A - \lambda I| = 0$

i.e. $\begin{vmatrix} 2-\lambda & 3 \\ 2 & 1-\lambda \end{vmatrix} = 0$ and $\lambda^2 - 3\lambda - 4 = 0$

$\Rightarrow \lambda_1 = -1$ and $\lambda_2 = 4$

To find the corresponding eigenvectors .. $\begin{pmatrix} 2-\lambda & 3 \\ 2 & 1-\lambda \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

$\lambda_1 = -1 \Rightarrow \begin{pmatrix} 3 & 3 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ i.e. $3x + 3y = 0$

i.e. $x/y = -1$

and $\underset{\sim}{x}_1 = \alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, α a scalar

$\lambda_2 = 4 \Rightarrow \begin{pmatrix} -2 & 3 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ i.e. $-2x + 3y = 0$

i.e. $x/y = 3/2$

and $\underset{\sim}{x}_2 = \beta \begin{pmatrix} 3 \\ 2 \end{pmatrix}$, β a scalar

Now return to the differential equation: $\frac{d}{dt}\underset{\sim}{x} = A\underset{\sim}{x} = \lambda \underset{\sim}{x}$

i.e. $\frac{d}{dt}\underset{\sim}{x} = \lambda \underset{\sim}{x}$ i.e. eigenvectors have a time-dependent amplitude of the form $e^{\lambda t}$.

Since it is a 2×2 system, the two eigenvectors form a complete set of linearly independent basis vectors for the solution space.

Thus, the general solution of system $\frac{d}{dt}\underset{\sim}{x} = A\underset{\sim}{x}$ can be written as a linear combination of these basis vectors

i.e. general solution is $\underset{\sim}{x} = c_1 e^{\lambda_1 t} \underset{\sim}{x}_1 + c_2 e^{\lambda_2 t} \underset{\sim}{x}_2$

i.e. $$\boxed{\underset{\sim}{x} = c_1 e^{-t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 e^{4t} \begin{pmatrix} 3 \\ 2 \end{pmatrix}}$$

Where c_1, c_2 are scalars, and any particular solution of the system

$$\frac{d}{dt}\underset{\sim}{x} = A\underset{\sim}{x}$$

is given by a particular choice of the constants c_1 and c_2.