AN OUTLINE OF "LINEAR ALGEBRA"

Prof. S.D.Agashe
Department of Electrical Engineering
Indian Institute of Technology Bombay
Phone: +91-22-2576-7411, email: eesdaia@ee.iitb.ac.in

Nov 24, 2008

1 Linear Algebra $\neq$ Matrix Theory

2

"Simultaneous" Linear Equations, or, Systems of Linear Equations – How do they "arise"?

2.1

Does the number of equations have to be equal to the number of unknowns or variables?

2.2

The set of **all** solutions of a system or the **solution set** of the system

- there may be **no** solution : the system is **inconsistent**, the equations **incompatible** with one another ; the system may be **over determined**
- there may be **at least one** solution, the system is **consistent**, but
  - the system may have **more than one** solution, may be **incomplete, underdetermined :** or
  - the system may have **one and only one** solution , the solution may be **unique** ; this is, of course, good, but may not happen.

2.3 Theorem of the Alternatives:

A system of linear equations may have **no** solution, **exactly one** solution, or **infinitely many** solutions. Obviously only **one** of the alternatives holds.

2.3.1

If the number of equations is **less** than the number of unknowns, **either** there is **no** solution, **or** there are infinitely many solutions.

3 How to solve a system $Ax = b$?

Here, b is a **column** , x is the **column** of the **variables** or **unknowns** , and A is the matrix of coefficients. Very rarely the system may be written as $yB = c$ where y and c are rows.

3.1

A may not be "square", but may be "rectangular " , "tall" or "broad" , so we cannot get away with

$$x = A^{-1}b \tag{1}$$

i.e.,find the inverse of the matrix, so, we have to go beyond inverse, determinants, etc.

3.2 Classical method of "Elimination of Variables":

This involves operations on the equations, changing them but not the variables or unknowns; when working with the matrix-vector form, $Ax = b$, this involves doing "elementary" operations on the rows of the matrix A, or row operations, this corresponds to choosing **any** equation, choosing **any** non zero coefficient occurring in it as the **pivot** element and eliminating the associated variable from **all** the **other** equations.

3.2.1

- It is not necessary that the (1,1) entry of the matrix has to be chosen as the **first** pivot element.
- The "elementary" operation of **interchanging** two rows is not **necessary**.
- The "elementary" operation of "scaling" a row is not **necessary**.
- The "elementary" operation of **changing** a row by adding a (non zero) multiple of another pivot row to it is **sufficient**.
- Since the **equations** are changed the right hand side numbers also get changed.

3.2.2

By using Row operations successively, the equations are changed into a form where

- **Inconsistency**, if any, becomes evident; the entries in a row of coefficients are zero but the corresponding number on the right hand side is not zero.
- **Consistency and completeness** : each unknown occurs in one and only equation and so can be solved "easily".
- **Consistency and Incompleteness**: some work is required to write the set of solutions in the form

  $$\underline{x} = \underline{v_0} + p_1\underline{v_1} + p_2\underline{v_2} + \ldots\ldots p_k\underline{v_k} \tag{2}$$

  where $\underline{v_0}, \underline{v_1}, \ldots \underline{v_k}$ are some columns of numbers and the "coefficients" $p_1, p_2, \ldots .p_k$ can be any arbitrary numbers ("independent free parameters"). The columns $\underline{v_1}, \underline{v_2}, \ldots .\underline{v_k}$ are "independent" solutions of the "corresponding" homogenous system

  $$A\underline{y} = \underline{0} \tag{3}$$

  and $\underline{v_0}$ is $\underline{a}$ "particular" solution of

  $$A\underline{x} = \underline{b} \tag{4}$$

  .

3.2.3

If $A\underline{x} = \underline{b}$ is to be solved for more than one column $\underline{b}$, or the entries of $\underline{b}$ are not specified, it is useful to calculate a "Row Operation" matrix at the same time as the row operations are being carried out on A. This is done by starting with an appropriate-sized identity matrix I, and working with the pair (A, I). when A becomes row-reduced to A_R, I will be changed to R, and it can be shown that

$$A_R = R.A \tag{5}$$

Instead of

$$A\underline{x} = \underline{b} \tag{6}$$

one can solve

$$A_R.\underline{x} = R.\underline{b} \tag{7}$$

3.3 "Column Operations" Method:

But there is **another** method of solving $A\underline{x} = \underline{b}$. It involves operating on **columns** of A, leaving $\underline{b}$ alone, that is, not operating on it. It is useful to keep track of the column operations by calculating a "Column Operation" matrix, starting with an appropriately sized identity matrix I . Thus I will be changed to a "Column Reduced" Matrix A_C while I is changed to C and

$$A_C = A.C \tag{8}$$

To solve

$$A\underline{x} = \underline{b} \tag{9}$$

one solves instead

$$A_C.\underline{y} = \underline{b} \tag{10}$$

and calculates x using

$$\underline{x} = C.\underline{y} \tag{11}$$

3.3.1

The method can thus be seen to involve a change of **variables** resulting also in a change of equations, but it is better not to worry about the **successive** change of variables; it is enough be aware that the "final " set of variables y is not the same as the "**original**" set of variables x.

3.3.2

In some of the transformed equations in new variables

$$A_C.\underline{y} = \underline{b} \tag{12}$$

only one **unknown** from the set y occurs so it can be solved "easily"; these calculated value are then substituted in the remaining equations, containing more than one unknown to check for consistency: we then obtain one particular solution.
Some columns of A_C may be "zero" so that the corresponding variables in y are free; using this fact and the relation

$$\underline{x} = C.\underline{y} \tag{13}$$

one can write a **general** or complete solution of

$$A\underline{x} = \underline{b} \tag{14}$$

as

$$\underline{x} = \underline{w_0} + p_1\underline{w_1} + p_2\underline{w_2} + \ldots..p_k\underline{w_k} \tag{15}$$

where $\underline{w_1}, \underline{w_2} \ldots\ldots \underline{w_k}$ are appropriate columns of the column operation matrix C.

3.3.3

Since one usually works with columns, the column operation method is more natural. Also, by using the column operation matrix, much more "information" can be easily obtained.

3.3.4

One can see that if the **solution set** of the homogenous system $A\underline{x} = \underline{0}$ is "non-empty", that is, if the system is consistent,it has the following two "properties " .

- If a column $\underline{v}$ is a solution so is p.$\underline{v}$ where p is any number;
- If the columns $\underline{v_1}, \underline{v_2}$ are two solutions, so is

$$p_1.\underline{v_1} + p_2.\underline{v_2} \tag{16}$$

  where p_1, p_2 are any two numbers;

- If $\underline{v_1}, \underline{v_2}, \ldots . \underline{v_k}$ are some solutions, so is any "linear combination"

$$p_1.\underline{v_1} + p_2.\underline{v_2} - + \ldots \ldots p_k.\underline{v_k} \tag{17}$$

  .

4 ENTER "LINEAR ALGEBRA" or "THEORY OF VECTOR SPACES"

4.1

The above observations on solution of systems of linear equations as well as some ideas from plane and solid geometry, and ideas from physics, **based** on these geometrical ideas, suggest the **abstraction** called "abstract **vector space.**"

A vector space $\underline{V}$ over a field $\underline{F}$ consists of a field F whose elements are referred to as "scalars" and a set $\underline{V}$, whose elements are referred to as "vectors" and a number of additional "things", namely:

- A binary operation $\bigoplus_V$ on V, called **addition** of vectors
- An operation $\bigodot_{F,V}$ involving an element of F and an element of V, producing an element of V. $\bigodot_{F,V}$ is called **"scalar" (-vector)** multiplication; **and**
- **assumptions** called **axioms**, that $V, F, \bigoplus_V and \bigodot_{F,V}$ satisfy or some properties that they have.

Homework: Look up a textbook for complete statement.

Note: A field itself involves abstract things like a set F and also some **operations**.

When we will talk about "vector spaces" ,we will "pretend" that we do not know anything about the elements of the sets V,F nor about the "operations" $\bigoplus_V, \bigodot_{F,V}$ except that certain properties hold or are satisfied.

4.2

One can, and should, of course think of **concrete vector spaces** side by side as a help in understanding or seeing the significance of abstract concepts. Some examples are :

- R,C as fields
- R^n_{col}, R^n_{row} as vector spaces over R

- The set $P_{\leq n}$ of all "polynomials" with degree $\leq n$, with "real" coefficients ;
- The set P of all polynomials ,with no restrictions on their degree ;
- $f : R \mapsto R$, the set of all real valued functions of real variables;
- The set of all "arrow" originating from a point and lying in a plane containing that point;
- The set of all "arrows" in "space " originating from a point.

4.3 SUBSPACE OF VECTOR SPACE :

The observation about the solution set of linear equations suggests the abstract concept of subspace.

A **subset W** of the set of vectors V of a **vector space** is called a subspace if

- It is "closed" under addition $\bigoplus_V$,
- It is "closed " under scalar multiplication $\bigodot_{F,V}$.

4.3.1

A subspace can also be "visualized" **geometrically**.

4.3.2

If W is a subspace of V, by using the operations from V, W can be made into a related or derived **vector space**. Such a derived vector space may have some properties which the "original" vector space from which it is derived does not have and "conversely".

4.3.3

Trivial subspaces $\{0_V\}$ and V itself.

4.4 SPAN OF A (FINITE) SET OF VECTORS $\{v_1, v_2, \ldots v_k\}$

Note: The v_i may not be the columns of numbers: In fact , we do not need to know what they are except that they belong to some abstract vector space V.
The 'span' is the set of all "linear" combinations of these vectors. These linear combination vectors are said to be "generated" by the vectors $\{v_1, v_2, \ldots \ldots v_k\}$.

4.4.1 SPAN OF A ARBITRARY SET W OF VECTORS FROM A VECTOR SPACE V :

It is the set of all **finite** linear combinations of elements of W.
Note: $\bigoplus_V$ is defined for **two** vectors, and can be extended for **three** , thousand, **billion**,... ... . i.e., any **finite** number of vectors , but **not** for any arbitrary "infinite number" of vectors.

4.5 INDEPENDENCE OF A SET OF VECTORS, AN "INDEPENDENT" SET OF VECTORS:

If not, **dependence** of a set of vectors, a **dependent** set of vectors, **dependence relation** between a set of vectors.

4.5.1

There is a more general notion of independence of a set of "conditions" or "properties" (and even of "Axioms"); dependence goes with redundancy, superfluity.

4.6 A BASIS B FOR A SUBSPACE W OF A VECTOR SPACE V

4.6.1

There is nothing like the basis for a subspace, or even for the vector space itself. However, $\{e_1, e_2, \ldots, e_n\}$ or $\{e_1^T, e_2^T, \ldots, e_n^T\}$ is a (sometimes useful) basis of "unit" vectors for R^n_{col}, R^n_{row}.

4.7 FINITE DIMENSIONAL VECTOR SPACE (OR SUBSPACE) AND THE DIMENSION OF SUCH A VECTOR SPACE OR SUBSPACE.

5 LINEAR FUNCTIONS AND TRANSFORMATIONS

Given two vector spaces V, W, over the same field F, with perhaps **different** operations $\bigoplus_V, \bigodot_{F,V}, \bigoplus_W$ and $\bigodot_{F,W}$ a function f on V into W denoted by $f : V \mapsto W$ is said to be **LINEAR** if

- It is homogenous, i.e,
$$f(\lambda \odot_{F,V} v) = \lambda \odot_{F,W} f(v), \tag{18}$$
and
- It is additive, i.e,
$$f(v_1 \oplus_V v_2) = [f(v_1)] \oplus_W [f(v_2)] \tag{19}$$

5.1

A $p \times q$ matrix A of real numbers could be viewed as being associated with, or producing, a linear function
$$A_{col} = R^q_{col} \mapsto R^p_{col}, \tag{20}$$
and another linear function

$$A_{row} : R^p_{row} \mapsto R^q_{row}$$

as follows

$$A_{col}(\underline{x}) = A\underline{x},$$
$$A_{row}(\underline{y}) = \underline{y}A,$$

here $\underline{x}$ is a column, $\underline{y}$ is a row.

Thus, a matrix can act on columns (of proper 'size') and also on rows (of proper size). The matrix may not be square,but could be rectangular.

5.1.1

Note: these definitions of associated functions do not involve any notion, and so, choice, of a " basis" for R^n_{col} or R^n_{row}.

5.2

However, given $f : V \mapsto W$, a linear function and two **ordered** bases, a basis B_V for V and basis B_W for W, if V is of dimension p and W of dimension q, we can associate with any abstract vector v in V, a **concrete** column vector $V^{B_V}_{col}$ *in* R^p_{col}, Similarly with a vector w in W.
Thus we have two representation functions,(or four), representing 'abstract' vectors as '**concrete'** columns (or rows).
Using this, we can associate with the function f a matrix $f^{col}_{B_V,B_W}$, such that, if $v \in V$ is represented by the column $V^{B_V}_{col}$, then $f(v) \in W$ will be represented by the column $[f(v)]^{B_W}_{col}$ and the matrix is defined

by

$$[f^{col}_{B_V,B_W}]\ v^{B_V}_{col} = [f(v)]^{B_W}_{col}. \tag{21}$$

Such a matrix can be "built up" column by column by choosing basis vectors in B_V one by one, "acting" on them by f, expanding the resulting vector with respect to the basis B_W, and writing the basis expansion coefficients as a column.
(If a vector v is shown or assumed to be a linear combination)

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots\ldots \lambda_p v_p \tag{22}$$

of the basis vector $v_1, v_2, \ldots..v_p$, such a sum is referred to as "basis expansion.")

5.2.1

Note: the matrix $f^{col}_{B_V,B_W}$, associated with or representing an abstract function f, depends on the choice B_V, B_W of bases for V and W.

5.2.2

Starting with a $p \times q$ matrix A, with no reference whatsoever to any bases(or even any vector space), we can think of A as giving rise to a linear **function**

$$A_{col} = R^q_{col} \mapsto R^p_{col}. \tag{23}$$

Now for R^q_{col} *and* R^p_{col}, one could choose bases B_V, B_W, not necessarily of "unit" vectors, and then obtain a matrix

$$(A_{col})^{col}_{B_V,B_W} \tag{24}$$

which will, in general, be quite different from A. Such a matrix is said to obtained from A by a "similarity transformation" or by a "change of basis" (it would be better to say: "by a choice of basis"). An important question(and problem) is: given a matrix A, can it be transformed, by a similarity transformation, into a "simpler matrix", e.g, identity matrix, diagonal matrix, companion form matrix, Jordan block form matrix etc.

6 KERNEL(OR NULL) AND IMAGE (OR RANGE)SUBSPACES ASSOCIATED WITH LINEAR FUNCTION $f : V \mapsto W$.

6.1

The solution set of a homogenous system

$$A\underline{x} = \underline{0}_{col} \tag{25}$$

suggests the concept of the kernel of f, $\underline{ker f}$.

6.2

The fact that $A\underline{x}$ can be seen to a linear combination of the columns of A, suggests the concept of the image of f, $\underline{imf.}$

6.3

The set of all linear combinations of the columns of a matrix is called the **column space** of the matrix A.

Similarly, **row space**.
The set of all solutions of the corresponding homogenous system

$$A\underline{x} = \underline{0} \tag{26}$$

is called the **column null space** of A. (Similarly, row null space).

Strang talks lovingly of the **four fundamental subspaces** associated with the matrix A;
however, with a linear function $f : V \mapsto W$. we have only **two** subspaces: $kerf$ is a subspace of V, imf is a subspace of W, so these are subspaces of the two different vector spaces V,W. This is because the elements of V and W are **abstract** vectors, not columns or rows, and f is an **abstract** linear function, not a matrix.
Given a $p \times q$ matrix A, if it is viewed as a linear function

$$A_{col} : R^q_{col} \mapsto R^p_{col}, \tag{27}$$

then its **column space** is a subspace of R^p_{col}, and its **column null space** is a subspace of R^q_{col}.

6.4

For a linear function $f : V \mapsto W$ $kerf$ may be a "trivial" subspace, either $\{0_v\}$ or V
$kerf = \{0_v\}$ if and only if f is "**one to one**".
$kerf = V$ if and only if f is the "**zero function**".

6.5

Similarly, imf may be trivial.
$imf = \{0_w\}$if and only if f is the **zero function**.
However, even if imf is one-to-one , imf may not be W. If $imf = W$, f is said to be **onto W**

6.6

If V is finite dimensional, so are $kerf$ and imf and

$$dim(kerf) + dim(imf) \;=\; dimV. \tag{28}$$

(W may or may not be finite dimensional)

6.7

Column operations are particularly useful here. If A, I are changed to A_c, C the non-zero columns of A_c (or even the corresponding columns of A), constitute a basis for the column space of A, and the columns of C, corresponding to the zero column of A, constitute a basis for the column null space.

Note: Column operations do not change the column space(or the image space) but may change the column null space. Column operation may change the row space but do not change the row-null space.

7 BREAKING UP INTO PARTS - TO REDUCE AND SIMPLIFY: DECOMPOSITION OF A VECTOR SPACE INTO SUM OF SUBSPACES

A vector space V is said to be **decomposed into two subspaces**, $W_1$and W_2, or said to be the **direct sum** of the subspaces $W_1$and W_2, if for each vector $v \in V$, there are two vectors w_1 and w_2, w_1 in W_1 and w_2 in W_2 such that:

$$v = w_1 + w_2 \tag{29}$$

and such vectors w_1, w_2 are unique, i.e., if there is also a decomposition of V given by

$$v = v_1 + v_2 \tag{30}$$

then v_1 must be equal to w_1 or w_2, and v_2 equal to w_2 and w_1 respectively.

One uses the symbolism

$$V = W_1 \oplus W_2. \tag{31}$$

Of course there is no use of this idea if the subspaces are trivial.

7.1

The definition can be extended to decomposition into more than two, but a finite number of subspaces.

7.2

If V is finite dimensional and

$$V = W_1 \oplus W_2 \tag{32}$$

then $W_1$and W_2 are also finite dimensional, and

$$dim(V) = dim(W_1) + dim(W_2). \tag{33}$$

So, a vector space V of dimension n can be decomposed into **at most** n subspaces, each of these being of dimension 1. This is finest possible decomposition.

7.3

The concept of decomposition for an **abstract** vector space is motivated by the idea of 'resolution' of vectors in geometry or physics(such as forces). One may refer to w_1 and w_2 as components of v.

The two subspaces W_1 and W_2 are 'smaller' than V and may have some special properties. But the real use of this concept is for the decomposition of linear functions.

7.4

Suppose $f : V \mapsto W$ is linear and V, W have decompositions such that

$$\begin{aligned} V &= V_1 \oplus V_2, \\ W &= W_1 \oplus W_2 \quad , \end{aligned}$$

and f takes elements of $V_1$to W_1, and elements of $V_2$to W_2.

Then, with f we can associate two functions :

$$\begin{aligned} f\mid_{V_1}\!: V_1 &\mapsto W_1 \quad , \\ f\mid_{V_2}\!: V_2 &\mapsto W_2. \end{aligned}$$

These are called the 'restrictions' of f to V_1 and V_2, respectively. The original function is nicely related to the two restrictions, which could be "simpler" than the original function. In particular, if one chooses bases $B_{V_1}$and $B_{V_2}$for V_1, V_2 and $B_{W_1}$and $B_{W_2}$for W_1, W_2, and bases $B_{V_1} \cup B_{V_2}$ for V, and $B_{W_1} \cup B_{W_2}$ for W, then the matrix representing f with respect to these two bases 'breaks up' into four parts, two of which are on the 'diagonal'. And the other two are zero matrices.

$$\begin{bmatrix} [\,] & 0 \\ 0 & [\,] \end{bmatrix}, \tag{34}$$

.

7.5

A very important special case is when the domain V of f and co-domain W of f are the same vector space, or $f : V \mapsto V$ is a **linear transformation** of the vector space V **into itself**. In this case to get a block diagonal decomposition of the representing matrix, it is necessary that the subspaces $V_1$and V_2 are "f - invariant", i.e.,f takes vectors in V_1 into vectors in $V_1$and vectors in V_2 into vectors into V_2

So we have the following problem: given $f : V \mapsto V$ find, if any, decomposition

$$V = V_1 \oplus V_2 \tag{35}$$

into f-invariant subspaces V_1 and V_2. One may also also want such a decomposition to be as fine as possible.

8 DECOMPOSITION USING ker/im SPACES

8.1

Given $f : V \mapsto V,\ V$ finite dimensional, it can be shown that

$$dimV = dim(im\ f) + dim(ker\ f) \tag{36}$$

Further it can be shown that the two subspaces $im\ f$ and $ker\ f$ are f-invariant. So, one would wish that:

$$V = im\ f \oplus ker\ f. \tag{37}$$

Unfortunately, this is so in exceptional cases.

8.1.1

If f is the zero function, $im\ f\ = \{0_v\}, ker\ f\ = V$, and so we have the trivial decomposition

$$V = \{0_v\} \oplus V. \tag{38}$$

Similarly, if f is a one-to-one transformation, then $im\ f = V,\ ker\ f = \{0_v\}$ and we get a trivial decomposition.

So, to get a non-trivial decomposition, f should be neither the zero function nor a one-to-one function, so that $ker\ f \neq \{0_v\}$ but, $ker\ f \supsetneqq \{0_v\}$ and $im\ f \neq V$ but, $imf \subsetneqq V$, i.e.,$ker\ f$ and $im\ f$ are non-trivial subspaces.

8.2

For a given $f : V \mapsto V$, one could find bases for $kerf$ and imf using column operations, and check whether these two bases put together give a bases for V. This can also be done using column operations. In fact, column operations can be used to check whether a given vector is some linear combination of some other given vectors, or whether a given finite set of vectors is independent

8.3

But if $V \neq kerf \oplus imf$, one need not despair. One can take the function $f \circ f$, also denoted by f^2, and called the **composition of f with itself**. (If f is represented by a matrix A, then f^2 is represented by the matrix $A \circ A$ or A^2. Thus, composition of linear functions and product of matrices are closely related - as also "addition" of linear functions, and "scalar multiple" of a linear function.) One can check whether

$$V = kerf^2 \oplus imf^2; \tag{39}$$

if this is not true, one can try f^3. It can be shown that if f is finite dimensional there is a **smallest** number $k,\ 1 \leq k \leq dimV$ such that

$$V = kerf^k \oplus imf^k. \tag{40}$$

Of course, this decomposition may turn out to be trivial, with $imf^k = \{0_V\},\ kerf^k = V$. In this case, f is said to be **nilpotent**, and k is said to be its **index of nilpotent**.

This method gives a decomposition, if it is non-trivial at all, into only **two**, subspaces. So, we look for other, "better" methods.

9 CYCLIC SUBSPACE AND MINIMUM POLYNOMIAL:

Given $f : V \mapsto V$, there is a simple way of generating f - invariant subspaces. Choose any non-zero vector $w \in V$, calculate f(w) and check if $f(w)$ depends on (and so , is a multiple of) w. If it does then

$$sp\{w\} \tag{41}$$

is f-invariant and of dimension one.

If f(w) is "independent" of w , then calculate $f(f(w))$ or $f^2(w)$ and check whether it is dependent on the independent set $\{w, f(w)\}$ containing two vectors. If it is, then

$$sp\{w, f(w)\} \tag{42}$$

is f-invariant and of dimension two.

So, it should be "clear" that given any non zero vector w of V and the linear function $f : V \mapsto V$ there is a smallest number $k,\ 1 \leq k \leq dimV$, such that the set

$$\{w, f(w), f^2(w), ...f^{\{k-1\}}(w)\} \tag{43}$$

is an independent set , and the vector f^k(w) is a linear combination of these vectors, and so the subspace

$$sp\{w, f(w), f^2(w), ...f^{\{k-1\}}(w)\} \tag{44}$$

is f-invariant. It is said to be the **cyclic subspace generated by w** under the action of f, or with respect to f. The set $\{w, f(w), f^{2(w)}, ...f^{\{k-1\}}(w)\}$ is said to be a **cyclic basis** for the cyclic subspace.

9.1

If f is - or is represented by - a matrix, and w is - or is represented by - a column, then using column operations one can determine such cyclic subspaces.

9.2

It may happen that there is a vector $w \in V$ such that, under the action of f, it generates V itself. In that case, V is said to be cyclic with respect to f, or even f is said to be cyclic.

9.2.1

If A is a special matrix, called a "companion" or "companion form" matrix:

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & \dots & -a_{n-1} & -a_n \end{bmatrix} \tag{45}$$

of size n, and we consider it as a function

$$A_{col} : R^n_{col} \mapsto R^n_{col} \tag{46}$$

then R^n_{col} is cyclic, or A is cyclic, with the unit vector e_n as a **cyclic generator**.

Note: There are some other forms of matrices which are also said to be "companion" or "companion form" by various authors.

9.3

A very important, and surprising, fact is the association of a **polynomial** with a vector $w \in V$, under the action of a given f, or with respect to a given f.

If

$$sp\{w, f(w), f^2(w), ...f^{\{k-1\}}(w)\}$$

is the **cyclic basis** for the cyclic subspace generated by w under the action of f, then $f^{(k)}(w)$ is a unique linear combination of these vectors, and so there is a **unique set** of numbers $a_{k-1}, a_{k-2}, \ldots, a_1, a_0$. such that

$$f^k(w) + a_{k-1}f^{k-1}(w) + \ldots + a_1 f(w) + a_0 w = 0_V. \tag{47}$$

The corresponding polynomial $p(s)$, given by

$$p(s) = p^k + a_{k-1}p^{k-1} + \ldots + a_1 p + a_0 \tag{48}$$

- or $p(\lambda)$, or $p(t)$, or even $p(x)$, if these symbols λ, t, x stand for a "dummy" variable - like s does, - this polynomial $p(s)$ is said to be the **minimum annihilating polynomial**, or **minimum polynomial**, **mp** for short, of the vector w under the action of f.

9.4

If A is an $n \times n$ matrix such that there is a column v_0 such that the mp of v_0 is of degree n, i.e., v_0 is a cyclic generator of R^n_{col}, with mp

$$p(s) = s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0, \tag{49}$$

then if the following **ordered bases** is chosen,

$$\{\underline{w_1}, \underline{w_2}, \underline{w_3} \ldots \underline{w_n}\}$$

where

$$\begin{aligned}
\underline{w_1} &= A^{n-1}\underline{w_0} + a_{n-1}A^{n-2}\underline{w_0} + \ldots + a_1\underline{w_0} \\
\underline{w_2} &= A^{n-2}\underline{w_0} + a_{n-1}A^{n-3}\underline{w_0} + \ldots + a_2\underline{w_0} \\
&\vdots \\
\underline{w_n} &= \underline{w_0},
\end{aligned}$$

the matrix representing A with respect to this bases will be the companion - form matrix mentioned in 9.2.1.

9.5 MINIMUM POLYNOMIAL OF A SUBSPACE AND OF THE WHOLE VECTOR SPACE:

The concept of mp of a **single** vector w with respect to a linear function f can be extended to the concept of mp of a set of vectors of a subspace and even of the whole vector space. It is simply the least degree monic(leading coefficient = 1) polynomial that, "annihilates"(makes into the zero vector) the subspace or the whole space. Such mp can be calculated as follows:

9.5.1 ALGORITHM FOR mp OF A FINITE-DIMENSIONAL SUBSPACE OR VECTOR SPACE:

Choose any bases for the subspace(or the whole space). Find out the mp of each of the bases vectors singly or separately. Then the mp of the subspace (or whole space) is the "LCM" (least common multiple) of the mp of the individual bases vectors.

9.5.2 HOW TO FIND THE LCM OF TWO POLYNOMIAL$\{p_1, p_2\}$?

NOT BY FACTORIZING EACH POLYNOMIAL (THE METHOD TAUGHT IN SCHOOL!) BUT BY FINDING OUT THE "GCD"(Greatest Common Divisor) or "HCF" (Highest Common Factor) OF THE POLYNOMIALS. THEN,

$$LCM = \frac{Product\ of\ polynomials}{GCD} \tag{50}$$

(There is no need to factorize when multiplying or dividing polynomials).

9.5.3 HOW TO FIND THE GCD OF TWO POLYNOMIALS WITHOUT FACTORIZING THEM?

Use the "EDA" (Euclidean Division Algorithm). The last non-zero remainder obtained in the EDA process is the GCD.

9.5.4 HOW TO FIND THE LCM OF A SET OF THREE OR MORE POLYNOMIALS $\{p_1, p_2, p_3, \ldots, p_k\}$.

First, find the LCM of p_1 and p_2, say p_{12}. Then find the LCM of p_{12} and p_3, say, p_{123}, and so on.

10 AN ALGORITHM FOR DECOMPOSITION THAT DOES NOT INVOLVE FACTORIZATION OF POLYNOMIALS

10.1

Why not factorization of polynomials(or determining their roots) when there are so many excellent computer programs available? Answer: The program gives only an **approximate** solution. There is no algorithm for finding roots of a real or complex - coefficient polynomial that involves only $+, -, \times$ and $\div$ operations!

10.2

1. First, determine the mp p_1 of the whole space.

2. Next, find out a vector w_1 whose mp is the mp of whole space. This is not as easy as it seems; we want an **algorithm**, no guesswork, choosing randomly, try this out, etc. There is an algorithm for finding out such a vector.(See my paper: "An algorithm for a result on minimal polynomials", **Linear Algebra and its applications**, vol. 357, 2002,pp. 291-293).

   In two special situations, both involving factorization, there is a method for obtaining such a vector that makes use of factorization.

   (a) If the mp is a "power" of an "irreducible" polynomial, say p^k, where p is an irreducible polynomial (no factors, e.g, $(s+2)$, or (s^2+s+1) - if we do not use **complex** numbers), then such a vector is found by choosing any bases for the subspace or whole space, and calculating their individual mp's. At least one of them will have mp $= p^k$.

   (b) If the mp is a product of two co-prime polynomials p_1 and p_2, say

   $$p = p_1 \circ p_2, \tag{51}$$

   then if (somehow) vectors v_1, v_2 are "known" whose mp's are p_1, p_2 , respectively, then the vector $(v_1 + v_2)$ has mp $= p$.

3. Next calculate the cyclic subspace V_1 generated by this special vector. If V_1 is the whole space (this will happen when the degree of the mp equals the dimension of the vector space), then stop. This algorithm will not give a decomposition. But if $V_1 \subsetneqq V$, we proceed further.

4. We calculate a new "object" - the **mp of the space relative to this subspace** V_1. This is an extension of the concept of mp of the space, not "relative" to anything - so that it now could be called the "absolute" mp - a new concept. (Actually, the original mp can be seen to be the mp relative to the trivial subspace $\{0_V\}$. To find the mp of a vector w relative to a given subspace V_1, we calculate $f(w), f^2(w), \ldots$ as before, except at each step, we check whether **some** linear combination of the vectors calculated up to any stage lies in the the subspace V_1. If it does, we stop. Thus the mp of w relative to V_1 will be a polynomial of **least** degree.

$$s^k + a_{k-1}s^{k-1} + \ldots + a_1 s + a_0 \tag{52}$$

if

(a) no linear combination $\{w, f(w), \ldots, f^{k-1}(w)\}$ is in V_1, and
(b) the l.c $f^k(w) + a_{k-1}f^{k-1}(w) + \ldots + a_1 f(w) + a_0 w$ is in V_1

The mp of the whole space relative to V_1 is, as before, the LCM of the mp's of vectors from any basis relative to V_1.

5. Next, find out a vector $w_2^{'}$ whose mp relative to V_1 is the same as the mp of whole space relative to V_1, say p_2. Then

$$p_2(f)(w_2^{'}) \in V_1 \tag{53}$$

and since V_1 is a cyclic subspace generated by our first vector w_1, there is a polynomial p_3 such that

$$p_2(f)(w_2^{'}) = p_3(f)w_1 \tag{54}$$

It has been shown that the polynomial p_2 divides p_3, i.e.,

$$p_3 = p_2 p_4$$
$$so \quad p_2(f)[w_2^{'} - p_4(f)(w_1)] = 0_V.$$

The vector $w_2^{'} - p_4(f)(w_1) = w_2$, say, is our next cyclic generator. The cyclic subspace V_2 generated by w_2 will be "independent" of the first cyclic subspace V_1.

Check whether

$V = V_1 \oplus V_2$.

If so, stop. No more decomposition is possible, using this algorithm.

If not find out the mp of the whole space relative to the subspace "$V_1 + V_2$" and proceed further.

If V is finite dimensional , this process will end after a finite number of steps.

6. The polynomials $p_1, p_2 \ldots$, obtained in this procedure are called the **invariant polynomials** of the function. Each polynomial in the sequence divides all of its "predecessors", and the product of the polynomials equals the "characterstic polynomial" of f. If appropriate bases are chosen for the cyclic subspaces $V_1, V_2, \ldots$, the matrix representing f will be a "block - diagonal" matrix, each diagonal block being a companion - form matrix. Such a matrix is known as the **FROBENIUS** or **RATIONAL CANONICAL FORM**.

11 ANOTHER ALGORITHM FOR DECOMPOSITION THAT DOES INVOLVE FACTORISATION OF POLYNOMIALS

11.1

Suppose the mp of the whole space p "has" a factorization

$$p = p_1 p_2 \tag{55}$$

where p_1 and p_2 are "coprime" (have no common factor, GCD is equal to constant polynomial), and no factorisation of p_1, p_2 is known. Then

$$\begin{aligned} ker(p_1(f)) &= im(p_2(f)) \\ ker(p_2(f)) &= im(p_1(f)) \\ V = ker(p_1(f)) &\oplus im(p_1(f)) \\ = ker(p_2(f)) &\oplus im(p_2(f)) \end{aligned}$$

Now apply the algorithm of the previous section 10 to these subspaces.

Example:

$$\begin{aligned} p(s) &= (s-1)^3(s^2+s+1)^2 \\ then \quad p_1(s) &= (s-1)^3,\ p_2(s) = (s^2+s+1)^2. \end{aligned}$$

11.2

Of course if p has a factorization into three or more piecewise co-prime polynomials, say

$$p = p_1 p_2 \dots p_k \tag{56}$$

it should be clear what is to be done.

Example:

$$p(s) = (s-1)^3(s+1)^2(s+3) \tag{57}$$

In this case, one will obtain the **JORDAN CANONICAL FORM.**