I LINEAR ALGEBRA

A. Fields.

A field is a set of elements in which a pair of operations called multiplication and addition is defined analogous to the operations of multiplication and addition in the real number system (which is itself an example of a field). In each field F there exist unique elements called o and 1 which, under the operations of addition and multiplication, behave with respect to all the other elements of F exactly as their correspondents in the real number system. In two respects, the analogy is not complete: 1)multiplication is not assumed to be commutative in every field, and 2)a field may have only a finite number of elements.

More exactly, a field is a set of elements which, under the above-mentioned operation of addition, forms an additive abelian group and for which the elements, exclusive of zero, form a multiplicative group and, finally, in which the two group operations are connected by the distributive law. Furthermore, the product of o and any element is defined to be o.

If multiplication in the field is commutative then the field is called a commutative field.

B. Vector Spaces.

If V is an additive abelian group with elements A,B,..., F a field with elements a,b,..., and if for each $a \varepsilon F$ and $A \varepsilon V$ the product aA denotes an element of V, then V is called a (left) vector space over F if the following assumptions hold:

1) $a(A + B) = aA + aB$

2) $(a + b) A = aA + bA$

3) $a(bA) = (ab) A$

4) $1A = A$

The reader may readily verify that if V is a vector space over F, then $oA = 0$ and $a0 = 0$ where o is the zero element of F and 0 that of V. For example, the first relation follows from the equations:

$$aA = (a + o) A = aA + oA$$

Sometimes products between elements of F and V are written in the form Aa in which case V is called a right vector space over F to distinguish it from the previous case where multiplication by field elements is from the left. If, in the discussion, left and right vector spaces do not occur simultaneously, we shall simply use the term "vector space."

C. Homogeneous Linear Equations.

If in a field F, a_{ij} $i=1,2,\ldots,m$, $j=1,2,\ldots,n$ are $m.n$ elements, it is frequently necessary to know conditions guaranteeing the existence of elements in F such that the following equations are satisfied:

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = o \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = o. \end{array} \tag{1}$$

The reader will recall that such equations are called linear homogeneous equations, and a set of elements, $x_1, x_2, \ldots, x_n$ of F, for which all the above equations are true, is called a solution of the system. If not all of the elements $x_1, x_2, \ldots, x_n$

are o the solution is called <u>non-trivial</u>; otherwise, it is called <u>trivial</u>.

<u>THEOREM 1</u>: <u>A system of linear homogeneous equations always has a non-trivial solution if the number of unknowns exceeds the number of equations.</u>

The proof of this follows the method familiar to most high school students, namely, successive elimination of unknowns. It is, of course, obvious that one homogeneous equation $a_1x_1 + a_2x_2 + \ldots + a_nx_n = o$ $n>1$, has a non-trivial solution. Indeed, if one of the a_i's is o, say $a_1 = o$, then $x_1 = 1$, $x_2 = x_3 = \ldots = x_n = o$ will serve as a solution; otherwise, $x_1 = a_2$, $x_2 = -a_1$ and $x_3 = x_4 = \ldots = x_n = o$ is a solution.

We shall proceed by complete induction. Let us suppose that each system of k equations in more than k unknowns has a non-trivial solution when $k<m$. In the system of equations (1) we assume that $n>m$, and denote the expression $a_{i1}x_1 + \ldots + a_{in}x_n$ by L_i, $i=1,2,\ldots,m$. We seek elements $x_1, \ldots, x_n$ not all o such that $L_1 = L_2 = \ldots = L_m = o$. If $a_{ij} = o$ for each i and j, then any choice of $x_1,\ldots, x_n$ will serve as a solution. If not all a_{ij} are o, then we may assume that $a_{11} \neq o$, for the order in which the equations are written or in which the unknowns are numbered has no influence on the existence or non-existence of a simultaneous solution. We can find a non-trivial solution to our given system of equations, if and only if we can find a non-trivial solution to the following system:

$$L_1 = o$$

$$L_2 - \frac{a_{21}}{a_{11}} L_1 = o$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \qquad \cdot$$

$$L_m - \frac{a_{m1}}{a_{11}} L_1 = o$$

For, if $x_1,\ldots,x_n$ is a solution of these latter equations then, since $L_1 = o$, the second term in each of the remaining equations is o and, hence, $L_2 = L_3 = \ldots = L_m = o$. Conversely, if (1) is satisfied, then the new system is clearly satisfied. The reader will notice that the new system was set up in such a way as to "eliminate" x_1 from the last n-1 equations. Furthermore, if a non-trivial solution of the last n-1 equations, when viewed as equations in $x_2,\ldots,x_n$, exists then taking $x_1 = -\frac{1}{a_{11}}(a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n)$ would give us a solution to the whole system. However, the last n-1 equations have a solution by our inductive assumption, from which the theorem follows.

Remark: If the linear homogeneous equations had been written in the form $\Sigma x_j a_{ij} = o$, $i=1,2,\ldots,n$, the above theorem would still hold and with the same proof although with the order in which terms are written changed in a few instances.

D. Dependence and Independence of Vectors.

In a vector space V over a field F, the vectors $A_1,\ldots,A_n$ are called dependent if there exist elements $x_1,\ldots,x_n$ not all o of F such that $x_1A_1 + x_2A_2 + \ldots + x_nA_n = 0$. If the vectors $A_1, \ldots, A_n$ are not dependent, they are called independent.

The dimension of a vector space V over a field F is the maximum number of independent elements in V. Thus, the dimension of V is n if there are n independent elements in V, but no set of more than n independent elements.

A system $A_1, \ldots, A_m$ of elements in V is called a generating system of V if each element A of V can be expressed linearly in terms of $A_1, \ldots, A_m$, i.e., $A = \sum_{i=1}^{m} a_i A_i$ for a suitable choice of a_i, $i=1, \ldots, m$, in F.

THEOREM 2: In any generating system the maximum number of independent vectors is equal to the dimension of the vector space.

Let $A_1, \ldots, A_m$ be a generating system of a vector space V of dimension n. Let r be the maximum number of independent elements in the generating system. By a suitable reordering of the generators we may assume $A_1, \ldots, A_r$ independent. By the definition of dimension, it follows that $r \leq n$. For each j, $A_1, \ldots, A_r$, A_{r+j} are dependent, and in the relation

$$a_1A_1 + a_2A_2 + \ldots + a_rA_r + a_{r+j}\,A_{r+j} = 0$$

expressing this, $a_{r+j} \neq 0$, for the contrary would assert the dependence of $A_1, \ldots, A_r$. Thus,

$$A_{r+j} = -\frac{1}{a_{r+j}}\left[a_1A_1 + a_2A_2 + \ldots + a_rA_r\right].$$

It follows that $A_1, \ldots, A_r$ is also a generating system since in the linear relation for any element of V the terms involving A_{r+j}, $j \neq 0$, can all be replaced by linear expressions in $A_1, \ldots, A_r$.

Now, let $B_1, \ldots, B_t$ be any system of vectors in V where $t > r$, then there exist a_{ij} such that $B_j = \sum_{i=1}^{r} a_{ij} A_i$, $j=1,2,\ldots,t$, since the A_i's form a generating system. If we can show that $B_1, \ldots, B_t$ are dependent this will give us $r \geq n$, and the theorem will follow from this together with the previous inequality $r \leq n$. Thus, we must exhibit the existence of a non-trivial solution out of F of the equation $x_1B_1 + x_2B_2 + \ldots + x_tB_t = 0$. To this end, it will be sufficient to choose the x_i's so as to satisfy the linear equations $\sum_{j=1}^{t} a_{ij} x_j = o$, $i=1,2,\ldots,r$, since these expressions will be the coefficients of A_i when in $\sum_{j=1}^{t} x_jB_j$ the B_j's are replaced by $\sum_{i=1}^{r} a_{ij} A_i$ and terms are collected. A solution to the equations $\sum_{j=1}^{t} a_{ij} x_j = o$, $i=1,2,\ldots,r$, always exists by Theorem 1.

Remark: Any n independent vectors $A_1,\ldots,A_n$ in an n dimensional vector space form a generating system. For any vector A, the vectors A, $A_1,\ldots, A_n$ are dependent and the coefficient of A, in the dependence relation, cannot be zero. Solving for A in terms of $A_1, \ldots, A_n$, exhibits $A_1, \ldots, A_n$ as a generating system.

A subset of a vector space is called a subspace if it is a subgroup of the vector space and if, in addition, the multiplication of any element in the subset by any element of the field is also in the subset. If $A_1,\ldots,A_s$ are elements of a vector space V, then the set of all elements of the form $a_1A_1 + \ldots + a_sA_s$ clearly forms a subspace of V. It is also evident, from the definition of dimension, that the dimension of any subspace never exceeds the dimension of the whole vector space.

An s-tuple of elements $(a_1, \ldots, a_s)$ in a field F will be called a <u>row vector</u>. The totality of such s-tuples form a vector space if we define

α) $(a_1, a_2, \ldots, a_s) = (b_1, b_2, \ldots b_s)$ if and only if $a_i = b_i$, $i = 1, \ldots, s$,

β) $(a_1, a_2, \ldots, a_s) + (b_1, b_2, \ldots, b_s) = (a_1 + b_1, a_2 + b_2, \ldots, a_s + b_s)$

γ) $b(a_1, a_2, \ldots, a_s) = (ba_1, ba_2, \ldots, ba_s)$, for b an element of F.

When the s-tuples are written vertically, $\begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_s \end{pmatrix}$ they will be called <u>column vectors</u>.

<u>THEOREM 3</u>. <u>The row (column) vector space F^n of all n-tuples from a field F is a vector space of dimension n over F.</u>

The n elements

$$\varepsilon_1 = (1, o, o, \ldots, o)$$
$$\varepsilon_2 = (o, 1, o, \ldots, o)$$
$$\vdots$$
$$\varepsilon_n = (o, o, \ldots, o, 1)$$

are independent and generate F^n. Both remarks follow from the relation $(a_1, a_2, \ldots, a_n) = \Sigma a_i \varepsilon_i$.

We call a rectangular array

$$\begin{pmatrix} a_{11} a_{12} \cdots a_{1n} \\ a_{21} a_{22} \cdots a_{2n} \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ a_{m1} a_{m2} \cdots a_{mn} \end{pmatrix}$$

of elements of a field F a <u>matrix</u>. By the <u>right row rank</u> of a matrix, we mean the maximum number of independent row vectors among the rows $(a_{i1},\dots,a_{in})$ of the matrix when multiplication by field elements is from the right. Similarly, we define left row rank, right column rank and left column rank.

THEOREM 4. In any matrix the right column rank equals the left row rank and the left column rank equals the right row rank. If the field is commutative, these four numbers are equal to each other and are called the rank of the matrix.

Call the column vectors of the matrix $C_1,\dots,C_n$ and the row vectors $R_1,\dots,R_m$. The column vector 0 is $\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ and any dependence $C_1x_1 + C_2x_2 + \dots + C_nx_n = 0$ is equivalent to a solution of the equations

$$
(1) \qquad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= 0 \\ \vdots \qquad\qquad\qquad & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= 0 . \end{aligned}
$$

Any change in the order in which the rows of the matrix are written gives rise to the same system of equations and, hence, does not change the column rank of the matrix, but also does not change the row rank since the changed matrix would have the same set of row vectors. Call c the right column rank and r the left row rank of the matrix. By the above remarks we may assume that the first r rows are independent row vectors. The row vector space generated by all the rows of the matrix has, by Theorem 1, the dimension r and is even generated by the first r rows. Thus, each row after the r^{th} is linearly expressible in terms of the first r rows. Consequently, any solution of the first r equations in (1) will be a solution of the entire system since any of the last n-r equations is obtainable as a linear combination of the first r. Conversely, any solution of (1) will also be a solution of the first r equations. This means that the matrix

$$\begin{pmatrix} a_{11} a_{12} \cdots a_{1n} \\ \vdots \quad \vdots \quad\quad \vdots \\ a_{r1} a_{r2} \cdots a_{rn} \end{pmatrix}$$

consisting of the first r rows of the original matrix has the same right column rank as the original. It has also the same left row rank since the r rows were chosen independent. But the column rank of the amputated matrix cannot exceed r by Theorem 3. Hence, $c \leq r$. Similarly, calling c' the left column rank and r' the right row rank, $c' \leq r'$. If we form the transpose of the original matrix, that is, replace rows by columns and columns by rows, then the left row rank of the transposed matrix equals the left column rank of the original. If then to the transposed matrix we apply the above considerations we arrive at $r \leq c$ and $r' \leq c'$.

E. Non-homogeneous Linear Equations.

The system of non-homogeneous linear equations

$$
(2) \qquad \begin{matrix}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\
a_{21}x_1 + \cdots\cdots\cdots + a_{2n}x_n = b_2 \\
\vdots \qquad\qquad\qquad \vdots \qquad \vdots \\
a_{m1}x_1 + \cdots\cdots\cdots + a_{mn}x_n = b_m
\end{matrix}
$$

has a solution if and only if the column vector $\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ lies in the space generated by the vectors $\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$

This means that there is a solution if and only if the right column rank of the matrix $\begin{pmatrix} a_{11} \cdots a_{1n} \\ \vdots \quad\;\; \vdots \\ a_{m1} \cdots a_{mn} \end{pmatrix}$ is the same as the right column rank of the augmented matrix $\begin{pmatrix} a_{11} \cdots a_{1n} b_1 \\ \vdots \quad\;\; \vdots \quad \vdots \\ a_{m1} \cdots a_{mn} b_m \end{pmatrix}$ since the vector space generated by the original must be the same as the vector space generated by the augmented matrix and in either case the dimension is the same as the rank of the matrix by Theorem 2.

By Theorem 4, this means that the row ranks are equal. Conversely, if the row rank of the augmented matrix is the same as the row rank of the original matrix, the column ranks will be the same and the equations will have a solution.

If the equations (2) have a solution, then any relation among the rows of the original matrix subsists among the rows of the augmented matrix. For equations (2) this merely means that like combinations of equals are equal. Conversely, if each relation which subsists between the rows of the original matrix also subsists between the rows of the augmented matrix, then the row rank of the augmented matrix is the same as the row rank of the original matrix. <u>In terms of the equations this means that there will exist a solution if and only if the equations are consistent, i.e., if and only if any dependence between the left hand sides of the equations also holds between the right sides.</u>

<u>THEOREM 5.</u> <u>If in equations (2) m = n, there exists a unique solution if and only if the corresponding homogeneous equations</u>

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = 0 \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = 0 \end{array}$$

<u>have only the trivial solution.</u>

If they have only the trivial solution, then the column vectors are independent. It follows that the original n equations in n unknowns will have a unique solution if they have any solution, since the difference, term by term, of two distinct solutions would be a non-trivial solution of the homogeneous equations. A solution would exist since the n independent column vectors form a generating system for the n-dimensional space of column vectors.

Conversely, let us suppose our equations have one and only one solution. In this case, the homogeneous equations added term by term to a solution

of the original equations would yield a new solutio
to the original equations. Hence the homogeneous
equations have only the trivial solution.