Lecture #1

Introduction

For those of you who have previously studied the material presented in these lectures, I hope it will revive and organize old memories. For those who have not seen much of the material, there is a risk that the course will be overwhelming, because a great deal will be presented in a short time. I will try to address the needs of both types of students by providing a structure that I hope will link the various topics tightly enough as to make it possible to retain them. You should also realize that to learn mathematics you need to do many practice problems. Mathematics is like a foreign language. It is learned not only through insight but by frequent repetition. I will assign daily problem sets. You should try to do these and then perhaps make up some more of your own to do.

An important thing to remember about mathematics is that the main purpose of its abstractions is to simplify problems and put you in a position to calculate in situations where you might otherwise be confused by complex impressions. The realization that an abstract concept can simplify life should make mathematics less daunting and easier to grasp and control.

Linear Algebra

1) Simultaneous Linear Equations

We start with a simple example of how one solves two simultaneous linear equations.

Example:

$$x_1 + 7x_2 = -57$$
$$12x_1 + 3x_2 = 45.$$

Subtract one third of the second equation from four times the first.

$$4x_1 + 28x_2 = -228$$
$$4x_1 + x_2 = 15$$
$$\overline{\qquad 27x_2 = -243}$$

It follows that $x_2 = -\dfrac{243}{27} = -9.$

Substituting this value for x_2 into the first equation, we obtain

$$x_1 = -7x_2 - 57 = 63 - 57 = 6.$$

Matrix Representation of Simultaneous Linear Equations

We need convenient notation to describe systematically the solution procedure just used. Write M simultaneous linear equations in N unknowns as follows:

$$a_{11}x_1 + a_{12}x_2 + ... + a_{1N}x_N = y_1$$
$$a_{21}x_1 + a_{22}x_2 + ... + a_{2N}x_N = y_2$$
$$\vdots$$
$$a_{M1}x_1 + a_{M2}x_2 + ... + a_{MN}x_N = y_M,$$

where the coefficients a_{mn} and y_m are numbers, and $x_1, x_2, ... , x_N$ are unknowns. These equations may be written as the single equation Ax = y, where x is the N-vector $(x_1, ... , x_N)$, y is the M-vector $(y_1, ... , y_M)$, and A is the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & . . . & a_{1N} \\ a_{21} & a_{22} & . . . & a_{2N} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_{M1} & & . . . . & a_{MN} \end{pmatrix}.$$

A is called an MxN matrix, which means that it has M rows and N columns. In writing the equation Ax = y, think of x as an Nx1 matrix and y as an Mx1 matrix, so that the equation may be visualized as

$$\begin{pmatrix} a_{11} & a_{12} & . . . & a_{1N} \\ a_{21} & a_{22} & . . . & a_{2N} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_{M1} & & . . . . & a_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_M \end{pmatrix} \qquad (1.1)$$

or as

$$
\begin{pmatrix} a_{11}x_1 + \dots + a_{1N}x_N \\ a_{21}x_1 + \dots + a_{2N}x_N \\ \cdot \\ \cdot \\ \cdot \\ a_{M1}x_1 + \dots + a_{MN}x_N \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_M \end{pmatrix}. \qquad (1.2)
$$

In going from system 1.1 to system 1.2, take the Nx1 matrix x, rotate it counterclockwise by 90 degrees, place it on each of the successive rows of the MxN matrix A, multiply superimposed entries, and add the products.

Elementary Row Operations and Row Equivalence of Matrices

Definition: The following operations on the MxN matrix A, are known as elementary row operations:

a) multiply a row of A by a non-zero number,
b) interchange two rows, and
c) replace a row by that row plus c times another row, where c is a non-zero number.

Definition: If the MxN matrix B is obtained from A by a finite sequence of such operations, then B and A are said to be row equivalent.

Lemma 1.1: If A and B are row equivalent, then the equations Ax = 0 and Bx = 0 have the same solutions.

Proof: It is sufficient to show that the equations Ax = 0 and Bx = 0 have the same solutions if B is obtained from A by one elementary row operation. This is clearly true for operations of types (a) and (b) above. Consider an operation of type (c). Let a_m be the m^{th} row of A. Suppose that row m is replaced by row m plus c times row k, where $c \neq 0$ and $k \neq m$. If Ax = 0, then $\sum_{n=1}^{N} a_{mn}x_n = 0$ and $\sum_{n=1}^{N} a_{kn}x_n = 0$, so that

$$
\sum_{n=1}^{N} (a_{mn} + ca_{kn})x_n = \sum_{n=1}^{N} a_{mn}x_n + c\sum_{n=1}^{N} a_{kn}x_n = 0,
$$

and hence Bx = 0. If Bx = 0, then $\sum_{n=1}^{N} a_{kn}x_n = 0$ and

$$
0 = \sum_{n=1}^{N} (a_{mn} + ca_{kn})x_n = \sum_{n=1}^{N} a_{mn}x_n + c\sum_{n=1}^{N} a_{kn}x_n = \sum_{n=1}^{N} a_{mn}x_n + 0,
$$

so that $\sum_{n=1}^{N} a_{mn} x_n = 0$ and hence $Ax = 0$. ■

We will soon see that elementary row operations may be used to find a matrix B row equivalent to any matrix A such that the solutions of $Bx = 0$ are obvious. A similar technique can be used to find solutions of the equation $Ax = y$.

If y is an M vector, let $(A \vdots y)$ be the $M\text{x}(N+1)$ matrix

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1N} & y_1 \\ a_{21} & a_{22} & \dots & a_{2N} & y_2 \\ \cdot & & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ a_{M1} & a_{M2} & \dots & a_{MN} & y_M \end{pmatrix}.$$

By an argument similar to the proof of lemma 1.1, if the $M\text{x}(N+1)$ matrix $(B \vdots z)$ is obtained from the $M\text{x}(N+1)$ matrix $(A \vdots y)$ by elementary row operations, then the equations $Bx = z$ and $Ax = y$ have the same solutions. For instance, if $Ax = y$, then $Ax - y = 0$ and so $(A \vdots y) \begin{pmatrix} x \\ \cdots \\ -1 \end{pmatrix} = 0$, where $\begin{pmatrix} x \\ \cdots \\ -1 \end{pmatrix}$ is the vector $(x_1, \dots, x_N, -1)$ written as an $(N+1)\text{x}1$ matrix. Hence

$(B \vdots z) \begin{pmatrix} x \\ \cdots \\ -1 \end{pmatrix} = 0$, and so $Bx = z$. Similarly $Ax = y$, if $Bx = z$.

Definition: A matrix B is said to be row reduced if

a) the first non-zero entry in any row is 1, and
b) each column that contains the first non-zero entry of some row has all its other entries equal to 0.

Definition: The first non-zero entry of a row is said to be the leading non-zero entry of that row.

Lemma 1.2: Elementary row operations may be used to reduce any MxN matrix A to a row reduced matrix B.

Proof: The proof is almost self-evident. ■

Example: The following matrix is row reduced.

$$\begin{pmatrix} 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Definition: A matrix is a row reduced echelon matrix if

a) it is row reduced,
b) any row of zeros lies below all non-zero rows, and
c) if the non-zero rows are rows 1 through r and the leading non-zero entry of row m is in column n_m, for m = 1, ... , r, then $n_1 < n_2 < .... < n_r$.

Example: The following matrix is a row reduced echelon matrix.

$$\begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Lemma 1.3: Every matrix is row equivalent to a row reduced echelon matrix.

Proof: Again the proof is almost self-evident. ■

Example: The following steps reduce the matrix on the left to row reduced echelon form via elementary row operations.

$$\begin{pmatrix} 3 & 2 & 1 \\ 6 & 4 & 2 \\ 6 & 8 & 5 \end{pmatrix} \to \begin{pmatrix} 3 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 4 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 2/3 & 1/3 \\ 0 & 0 & 0 \\ 0 & 1 & 3/4 \end{pmatrix} \to \begin{pmatrix} 1 & 2/3 & 1/3 \\ 0 & 1 & 3/4 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1/6 \\ 0 & 1 & 3/4 \\ 0 & 0 & 0 \end{pmatrix}.$$

Remark: If B is a row reduced echelon matrix, the solutions of the equation Bx = 0 are obvious.

Example: The matrix $B = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ is in row reduced echelon form. One sees immediately that $B\begin{pmatrix} -3 \\ 1 \\ 1 \end{pmatrix} = 0$.

Theorem 1.4: If A is an MxN matrix such that M < N, then the equation Ax = 0 has a non-zero solution.

Proof: Let B be a row reduced echelon matrix that is row equivalent to A. By lemma 1.1, Ax = 0 if and only if Bx = 0. Let r be the number of non-zero rows of B, so that the non-zero rows of B are rows 1 through r. For m = 1, ... , r, let the first non-zero entry of row m be in column n_m, so that $n_1 < n_2 < .... < n_r$. Since $r \leq M < N$, there is an integer n such that $1 \leq n \leq N$ and $n \neq n_m$, for any m. Choose one such n, and let $x_n = 1$. For k such that k = 1, ... , N, let $x_k = 0$, if $k \neq n$ and if $k \neq n_m$, for m = 1, ..., r, and let $x_{n_m} = -b_{mn}$, if m = 1, ... , r. Since $b_{m,n_m} = 1$, $\sum_{k=1}^{N} b_{mk} x_k = b_{m,n_m} x_{n_m} + b_{mn} x_n = 1(-b_{mn}) + b_{mn}(1) = 0$, for m = 1, ... , r. Since the first r rows of B are its non-zero rows, it follows that Bx = 0 and hence Ax = 0. Since $x_n = 1$, $x \neq 0$. ■

Corollary 1.5: If A is an NxN matrix and if Ax = 0 has no non-zero solution, then A is row equivalent to the NxN identity matrix

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

Proof: Let B be a row reduced echelon matrix that is row equivalent to A. Since Ax = 0 has no non-zero solution, the equation Bx = 0 has no non-zero solution. It follows that the number of non-zero rows of B is N, for suppose that B has at least one row of zeros. Then the last row of B is zero. The first N–1 rows of B form an (N–1)xN matrix, call it C. By theorem 1.4, there is a non-zero N-vector x such that Cx = 0. Since the last row of B consists of zeros, Bx = 0 and we have a contradiction. Because the number of non-zero rows of B is N, B = I. ■

2) Vector Spaces

Consider the set $R^N = \{v = (v_1, ... , v_N) \mid v_n$ is a number, for all n$\}$, where N is a positive integer. We can define the following operations on vectors in R^N: if v and w belong to R^N, then $v + w = (v_1 + w_1, ... , v_N + w_N)$ and if c is a number and v belongs to R^N, then $cv = (cv_1, ... , cv_N)$. The following rules apply to these operations:

1) $v + w \in R^N$, if $v \in R^N$ and $w \in R^N$,

2) $cv \in R^N$, if c is a number and $v \in R^N$,
3) $v + w = w + v$, for all v and w in R^N,
4) there is $0 \in R^N$ such that $0 + v = v$, for all $v \in R^N$,
5) for all $v \in R^N$, there is a unique $-v \in R^N$ such that $v + (-v) = 0$,
6) $1v = v$, for all $v \in R^N$,
7) $(c_1 c_2)v = c_1(c_2 v)$, for all $v \in R^N$ and all numbers c_1 and c_2,
8) $(v + w) + z = v + (w + z)$, for all v, w, and z in R^N,
9) $c(v + w) = cv + cw$, for all numbers c and for all v and w in R^N,
10) $(c_1 + c_2)v = c_1 v + c_2 v$, for all numbers c_1 and c_2 and for all $v \in R^N$.

Definition: A vector space consists of a set V together with operations of addition and multiplication by numbers, denoted by $v + w$ and cv, for v and w in V and for all numbers c. These operations satisfy rules (1) - (10) above with R^N everywhere replaced by V.

R^N is an example of a vector space.

Definition: W is a subspace of V, if W is a subset of V and W is itself a vector space under the operations of addition and multiplication by numbers defined on V.

Remark: If W is a subset of a vector space V, then W is a subspace of V if $v + w \in W$ and $cv \in W$, for all vectors v and w in W and for all numbers c.

Example: $\{(v_1, v_2) \in R^2 \mid v_1 + v_2 = 0\}$ is a subspace of R^2.

Linear Independence and Bases

Definition: If V is a vector space, the vector v in V is said to be a linear combination of the vectors $v_1, \ldots, v_N$ in V if there are numbers $c_1, \ldots, c_N$ such that $v = c_1 v_1 + \ldots + c_N v_N$.

Definition: If $v_1, \ldots, v_N$ belong to the vector space V, their linear span is the set of all linear combinations of $v_1, \ldots, v_N$. The vectors $v_1, \ldots, v_N$ are said to span V, if V is the linear span of $v_1, \ldots, v_N$.

The linear span of $v_1, \ldots, v_N$ is a subspace of V and is the smallest subspace containing $v_1, \ldots, v_N$.

Example: R^2 is the linear span of the vectors (1, 0) and (0, 1).

Definitions: The vectors $v_1, \ldots, v_N$ in a vector space V are linearly dependent if there exist numbers $c_1, \ldots, c_N$, not all of which are zero, such that $c_1 v_1 + \ldots + c_N v_N = 0$.

The vectors $v_1, \ldots, v_N$ are linearly independent if they are not linearly dependent.

Example: The vectors (1, 0, 0), (0, 1, 0), and (1, 1, 0) are linearly dependent in R^3, since (1, 0, 0) + (0, 1, 0) − (1, 1, 0) = 0. The vectors (1, 0, 0), (0, 1, 0), and (0, 0, 1) are linearly independent, since the equations (0, 0, 0) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) = (a, b, c) imply that a = b = c = 0.

Definition: A basis for a vector space V is a set of linearly independent vectors in V that spans V.

Example: Let $e_n = (0, \ldots, 0, 1, 0, \ldots, 0) \in R^N$, where the 1 is in the n^{th} position. Then $e_1, \ldots, e_N$ is the standard basis of R^N.

Theorem 1.6: If $v_1, \ldots, v_M$ span a vector space V, then any independent set of vectors in V has no more than M elements.

Proof: I must show that if $w_1, \ldots, w_N$ are in V, where N > M, then $w_1, \ldots, w_N$ are linearly dependent. Since $v_1, \ldots, v_M$ span V, $w_n = \sum_{m=1}^{M} a_{mn} v_m$, for all n and for some numbers $a_{1n}, \ldots, a_{Mn}$. If $x_1, \ldots, x_N$ are numbers, then

$$\sum_{n=1}^{N} x_n w_n = \sum_{n=1}^{N} x_n \sum_{m=1}^{M} a_{mn} v_m = \sum_{n=1}^{N} \sum_{m=1}^{M} a_{mn} x_n v_m = \sum_{m=1}^{M} \left(\sum_{n=1}^{N} a_{mn} x_n \right) v_m .$$

Since M < N, theorem 1.4 implies that there exist numbers $x_1, \ldots, x_N$, not all zero, such that $\sum_{n=1}^{N} a_{mn} x_n = 0$, for m = 1, ... , M. Hence $x_1 w_1 + \ldots + x_N w_N = 0$ and so $w_1, \ldots, w_N$ are linearly dependent. ■

Definition: A vector space is finite dimensional, if it has a finite basis.

Corollary 1.7: If V is a finite dimensional vector space, then any two bases have the same number of elements.

Proof: If $v_1, \ldots, v_M$ and $w_1, \ldots, w_N$ are bases of V, then because $v_1, \ldots, v_M$ span V and $w_1, \ldots, w_N$ are independent, it follows that $N \le M$. Similarly, $M \le N$, so that N = M. ■

Definition: The dimension of V is the number of vectors in a basis of V. The dimension of V is denoted dim V.

Corollary 1.8: If V is a vector space of dimension N, then any N vectors in V that span V are linearly independent and so are a basis for V.

Proof: If $v_1, \ldots, v_N$ span V and are not linearly independent, then for some numbers $c_1, \ldots, c_N$, not all of which are zero, $c_1v_1 + \ldots + c_Nv_n = 0$. Without loss of generality, we may assume that $c_1 \neq 0$, so that $v_1 = -c_1^{-1}(c_2v_2 + \ldots + c_Nv_N)$. Since v_1 is a linear combination of $v_2, \ldots, v_N$, these N – 1 vectors span V. By theorem 1.6, there can be no more than N – 1 linearly independent vectors in V, so that dim V ≤ N – 1, contrary to the hypothesis that dim V = N. ■

A similar argument proves the following assertion.

Corollary 1.9: If V is a finite dimensional vector space, a basis for V is any smallest or minimal set of vectors in V that span V.

Lemma 1.10: If $v_1, \ldots, v_M$ are linearly independent vectors in V and $w \in V$ does not belong to the span of $v_1, \ldots, v_M$, then $v_1, \ldots, v_M$, w are linearly independent.

Proof: Suppose that $c_1v_1 + \ldots + c_Mv_M + bw = 0$. If $b \neq 0$, then $w = -b^{-1}c_1v_1 - \ldots - b^{-1}c_Mv_M$, which belongs to the span of $v_1, \ldots, v_M$. Since w does not belong to the span of $v_1, \ldots, v_M$, b must equal 0. Therefore $c_1v_1 + \ldots + c_Mv_M = 0$. Since $v_1, \ldots, v_M$ are linearly independent, $c_m = 0$, for all m, and so $v_1, \ldots, v_M$, w are linearly independent. ■

Corollary 1.11: If V is a vector space of dimension N, then any N linearly independent vectors in V span V and so are a basis for V.

Proof: If $v_1, \ldots, v_N$ are linearly independent and do not span V, then there is $v \in V$ that does not belong the linear span of $v_1, \ldots, v_N$, so that $v_1, \ldots, v_N$, v are linearly independent. Theorem 1.6 implies that this is impossible, since V has a basis of N vectors. ■

Theorem 1.12: If $v_1, \ldots, v_N$ is a basis for V and $v \in V$, then the numbers $c_1, \ldots, c_N$ such that $v = \sum_{n=1}^{N} c_nv_n$ are unique.

Proof: If $\sum_{n=1}^{N} c_nv_n = v = \sum_{n=1}^{N} a_nv_n$, then $\sum_{n=1}^{N} (c_n - a_n)v_n = 0$. Since $v_1, \ldots, v_N$ are independent, it follows that $c_n - a_n = 0$, for all n. ■

Theorem 1.13: If the vectors $v_1, \ldots., v_N$ span the vector space V and $0 < \dim V < \infty$, then some subset of $v_1, \ldots., v_N$ forms a basis for V.

Proof: Since dim V > 0, $v_n \neq 0$, for some n. Without loss of generality, we may assume that $v_1 \neq 0$. The vector v_1 by itself is a linearly independent set of vectors, so that there is a subset of $v_1, \ldots., v_N$ that is linearly independent. Therefore there is a largest subset of $v_1, \ldots., v_N$ that is linearly independent. Without loss of generality, we may let this subset be $v_1, \ldots., v_K$, where $K \leq N$. If K = N, then $v_1, \ldots., v_N$ is a basis for V. If K < N and n > K, then v_n must be a linear combination of $v_1, \ldots., v_K$, for otherwise by lemma 1.10, the set $v_1, \ldots., v_K, v_n$ would be linearly independent and would be larger than $v_1, \ldots., v_K$, which is impossible. Since $v_1, \ldots., v_N$ spans V, the set $v_1, \ldots., v_K$ spans V and so forms a basis for V. ■

A similar argument proves the following.

Theorem 1.14: If V is a finite dimensional, non-zero vector space, any largest or maximal set of linearly independent vectors in V is a basis for V.

This theorem suggests a way to construct a basis for a non-zero vector space V. Choose a non-zero vector v_1 in V. If v_1 spans V, it is a basis for V. We continue by induction on the number n of vectors chosen to belong to the basis. Suppose that the linearly independent vectors $v_1, \ldots., v_n$ have been chosen from V. If they span V, they form a basis. If not, choose a non-zero vector v_{n+1} from V that does not belong to the span of $v_1, \ldots., v_n$. By lemma 1.10, $v_1, \ldots., v_n, v_{n+1}$ are linearly independent. If V is finite dimensional, theorem 1.6 implies that this process must eventually end. That is, for large enough N, the linearly independent vectors $v_1, \ldots., v_N$ must span V and so form a basis for it.

Suppose that the dimension of the vector space V is N and we have K linearly independent vectors $v_1, \ldots., v_K$ in V, where K < N. The inductive process described in the previous paragraph may be used to construct a basis $v_1, \ldots., v_K, v_{K+1}, \ldots., v_N$ for V. The new basis is called an extension of $v_1, \ldots., v_K$ to a basis for V.

Theorem 1.15 Let W be a non-zero subspace of a finite dimensional vector space V such that $W \neq V$. Then W has a finite basis and dim W < dim V.

Proof: Since W is not zero, there is a non-zero vector w_1 in W. Because w_1 is non-zero, it is independent. From the inductive process described above, we know there is a sequence

$w_1, w_2, \ldots$ of independent vectors in W. Since these vectors are in V, which is finite dimensional, we know from theorem 1.6 that this sequence has no more than dim V members. Let $w_1, w_2, \ldots, w_M$ be a maximal such sequence, where $M \leq \dim V$. By the argument used in the previous two paragraphs, $w_1, w_2, \ldots, w_M$ span W and so form a basis for W, and so M = dim W. Since $W \neq V$, there is a non-zero vector v in V such that v does not belong to W. By lemma 1.10, $w_1, w_2, \ldots, w_M$, v are independent and so by theorem 1.14, $\dim V \geq M+1 > M = \dim W$. ■

The Row and Column ranks of a Matrix

If A is an MxN matrix, its rows are N-vectors and so belong to R^N and its columns are M-vectors and so belong to R^M. The linear span of the rows of A is a subspace of R^N called the row space of A, and the linear span of the columns of A is a subspace of R^M called the column space of A.

Definition: If A is an MxN matrix, the column rank of A is the dimension of the column space of A and the row rank of A is the dimension of the row space of A.

The objective of this section is to show that the column and row ranks of any matrix are equal. I prove this assertion by reducing the matrix to row reduced echelon form.

Lemma 1.16: If the MxN matrices A and B are row equivalent, then their row spaces are the equal.

Proof: It is sufficient to show that each of the elementary row operations does not change the row space. Clearly neither interchanging two rows nor multiplying a row by a non-zero number changes the row space. Consider replacing a row by the sum of that row and a multiple of another row. Without loss of generality, suppose that B is the matrix A with the first row of A replaced by that row plus d times the second row, where $d \neq 0$. If v belongs to the span of the rows of A, then $v = \sum_{m=1}^{M} c_m a_m$, where a_m is the m^{th} row of A and $c_1, \ldots, c_M$ are numbers. Then $v = c_1(a_1 + da_2) + (c_2 - c_1 d) a_2 + \sum_{m=3}^{M} c_m a_m = c_1 b_1 + (c_2 - c_1 d) b_2 + \sum_{m=3}^{M} c_m b_m$, where b_m is the m^{th} row of B. Therefore v belongs to the span of the rows of B. Similarly if v belongs to the span of the rows of B, $v = \sum_{m=1}^{M} c_m b_m$, for some numbers $c_1, \ldots, c_M$, and so $v = c_1 a_1 + (c_1 d + c_2) a_2 + \sum_{m=3}^{M} c_m a_m$ and hence lies in the span of the rows of A. This proves that the row spaces of A and B are the same. ■

The preceding lemma implies that if B is obtained from A by elementary operations, then the row ranks of A and B are the same.

Lemma 1.17: If the MxN matrices A and B are row equivalent, then their column ranks are the equal.

Proof: Let K be the column rank of A. By theorem 1.13, we may assume that K columns of A form a basis for its column space. Without loss of generality, we may assume that this basis consists of the first K columns of A. If $K < n \leq N$, then the n^{th} column of A is a linear combination of the first K columns. That is, $a^n = \sum_{k=1}^{K} c_k a^k$, where a^k denotes the k^{th} column of A. Therefore $Ax = 0$, where $x_k = c_k$, for $k = 1, \ldots, K$, $x_n = -1$, and $x_i = 0$, for $i > K$ and $i \neq n$. By lemma 1.1, $Bx = 0$ and therefore $b^n = \sum_{k=1}^{K} c_k b^k$, where b^k denotes the k^{th} column of B. Therefore the first K columns of B span the column space of B. I complete the proof by showing that the first K columns of B are independent. If they are not, then there exist numbers $c_1, \ldots, c_K$, not all of which are zero, such that $\sum_{k=1}^{K} c_k b^k = 0$. Let x be the N-vector defined by $x_k = c_k$, if $k = 1, \ldots, K$, and $x_n = 0$, if $n = K + 1, \ldots, N$. Then $Bx = 0$, so that by lemma 1.1, $Ax = 0$ and hence $\sum_{k=1}^{K} c_k a^k = 0$ and so the first K columns of A are not independent, contrary to hypothesis. Therefore the first K columns of B are linearly independent and so form a basis for the column space of B. Hence the column rank of B is K. ■

Theorem 1.18: The row and column ranks of any matrix are equal.

Proof: Let A be an MxN matrix and let B be a row reduced echelon matrix that is row equivalent to A. By lemmas 1.16 and 1.17, the row rank of A equals that of B and the column rank of A equals that of B, so that it is sufficient to show that the row and column ranks of B are equal. Let K be the number of non-zero rows of B, these being the first K rows, since B is a row reduced echelon matrix. I show that both the row and column ranks of B equal K.

For $k = 1, \ldots, K$, let the leading non-zero entry of the k^{th} row be in column n_k, where $n_1 < n_2 < \ldots < n_K$. Since this leading non-zero entry is the only non-zero entry of column n_k, the first K rows of B are linearly independent. Since the first K rows of B contain its only non-zero rows, they span the row space of B. Therefore the row rank of B equals K.

Similarly columns $n_1, n_2, \ldots, n_K$ of B are linearly independent. Since for $k = 1, \ldots, K$, $b_{k,n_k} = 1$ and $b_{m,n_k} = 0$, if $m \neq k$, it follows that if $n \neq n_k$, for any $k = 1, \ldots, K$, then $b^n = \sum_{k=1}^{K} b_{kn} b^{n_k}$, where b^n denotes the n^{th} column of B. Therefore columns $n_1, n_2, \ldots, n_K$ of B span the column space of B and so form a basis of that space. Hence the column rank of B equals K. ■