AN ABSTRACT OF THE THESIS OF

Barry Dow Biegun for the M.S. in Mathematics
(Degree) (Major)

Date thesis is presented April 5, 1965

Title EQUIVALENCE RELATIONS ON MATRICES

Abstract approved (Major professor)

Row equivalence, equivalence, and similarity of matrices are studied; some problems concerning an extension of these relations to infinite matrices are discussed.
EQUIVALENCE RELATIONS ON MATRICES

by

BARRY DOW BIEGUN

A THESIS
submitted to
OREGON STATE UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

MASTER OF SCIENCE

June 1965
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>I. Elementary Matrices</td>
<td>2</td>
</tr>
<tr>
<td>II. Row Equivalence</td>
<td>9</td>
</tr>
<tr>
<td>III. Equivalence</td>
<td>21</td>
</tr>
<tr>
<td>IV. Similarity</td>
<td>28</td>
</tr>
<tr>
<td>V. Infinite Matrices</td>
<td>35</td>
</tr>
<tr>
<td>Bibliography</td>
<td>44</td>
</tr>
</tbody>
</table>
EQUIVALENCE RELATIONS ON MATRICES

INTRODUCTION

It is our purpose to discuss certain well-known equivalence relations on matrices. Our work is based on Chapter 6 of Finkbeiner's *Introduction to Matrices and Linear Transformations* (1960). Our theorems correspond to the theorems of the above chapter, where proofs often appear greatly condensed or are left to the reader. We supplement the theorems by several lemmas. Whenever a specific previous result is assumed, it is accompanied by a page reference.

In the last section we discuss problems related to an extension of the results of Sections I - IV to infinite matrices.
I. ELEMENTARY MATRICES

Presently we wish to recall two systems which give rise to a matrix. In the notation of Finkbeiner (1960), let $V_m$ be a vector space with a fixed basis $\{\alpha_1, \ldots, \alpha_m\}$. Let $T$ be a linear transformation of $V_m$ into a vector space $W_n$, with the basis $\{\beta_1, \ldots, \beta_n\}$. For each $i=1, \ldots, m$, the image $\alpha_i T$ is a uniquely determined vector of $W_n$. Hence $\alpha_i T$ is uniquely determined as a linear combination of the $\beta$-vectors,

$$\alpha_i T = \sum_{j=1}^{n} a_{ij} \beta_j, \, i = 1, \ldots, m.$$  \hspace{1cm} (1)

With respect to the two bases, $T$ is completely determined by the $mn$ scalars $a_{ij}$. If we agree to a fixed ordering of the scalars, we can represent $T$ by the matrix

$$A = (a_{ij}), \, i=1, \ldots, m; j=1, \ldots, n.$$  Sometimes we will denote $A$ more compactly by $(a_{ij})_{mn}$.

In the system of $m$ linear equations in $n$ unknowns,

$$\sum_{j=1}^{n} a_{ij} x_j = c_i, \, i=1, \ldots, m,$$  \hspace{1cm} (2)

the array of coefficients $x_j$ also is uniquely represented by matrix $A$. The familiar elementary operations on the equations of (2) which reduce it to an equivalent system, have an analogue on the rows of a matrix.

**Definition.** The following are **elementary row operations** on a matrix:
(i) interchanging any two rows;
(ii) multiplying a row by a non-zero scalar;
(iii) adding one row to another.

Each operation above may be represented by a special matrix:

**Definition.** An $m \times m$ matrix obtained from the $m \times m$ identity matrix $I$ by performing a single elementary row operation is called an **elementary matrix**.

**Notation.** Recall that the $m \times m$ identity matrix $I = (\delta_{ij})$ where $\delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$. Let $E_{rs}$ denote a square matrix with $e_{rs} = 1$ and $e_{ij} = 0$ if $i \neq r$ or $j \neq s$. Thus $E_{rs}$ is a matrix with exactly one non-zero element. We define the elementary matrices in terms of $I$ and $E_{rs}$. Let $P_{ij}$ denote the elementary matrix obtained by interchanging the $i^{th}$ and $j^{th}$ rows of $I$. Then

$$P_{ij} = I - E_{ii} + E_{ji} - E_{jj} + E_{ij}.$$ 

Let $M_i(c)$ denote the elementary matrix obtained by multiplying the $i^{th}$ row of $I$ by a non-zero scalar $c$. Then

$$M_i(c) = I - E_{ii} + cE_{ii} = I + (c-1)E_{ii}.$$ 

Let $A_{ij}$ denote the elementary matrix obtained by adding the $i^{th}$ row of $I$ to its $j^{th}$ row. Then $A_{ij} = I + E_{ji}$.

**Lemma 1.** For any $m \times m$ matrix $A$, $IA = AI = A$.

**Proof.** The entry in the $(i,k)$ position of $IA$ is

$$\sum_{j=1}^{m} \delta_{ij} a_{jk}.$$ 

Since $\delta_{ij} = 0$ for all $i \neq j$, all the
terms in this sum are zero except $\delta_{ii}a_{ik} = a_{ik}$. Now $a_{ik}$ is the corresponding entry of $A$, so that $IA = A$. Similarly for $AI$, the entry of which in the $(i,k)$ position is 
$$\sum_{j=1}^{m} a_{ij} \delta_{jk} = a_{ik},$$
the corresponding entry of $AI$. Hence $AI = A$.

**Theorem 1.** Any elementary row operation can be performed on an $m \times n$ matrix $A$ by premultiplying $A$ by the corresponding $m \times m$ elementary matrix.

**Proof.** (i). $P_{ij}A = (I - E_{ii} + E_{ji} - E_{jj} + E_{ij})A$

$$= A - E_{ii}A + E_{ji}A - E_{jj}A + E_{ij}A,$$

since $IA = A$.

Consider $E_{ii}A$. The entry in the $(i,k)$ position of $E_{ii}A$ is 
$$\sum_{j=1}^{m} e_{ij}a_{jk} = 0 \text{ if } j \neq i. \text{ If } j = i, \sum_{i=1}^{1} e_{ii}a_{ik} = a_{ik}$$
since the other terms in this sum are zero. As $k$ ranges from 1 to $n$, $a_{ik}$ ranges from $a_{i1}$ to $a_{in}$. Hence the entry in the $i$th row of $E_{ii}A$ is the same as the corresponding entry in the $i$th row of $A$. Thus $E_{ii}A = A_{ii}$, a matrix with the $i$th row corresponding to the $i$th row of $A$ and zeros elsewhere. Next consider $E_{ji}A$. The entry in the $(j,k)$ position of $E_{ji}A$ is 
$$\sum_{i=1}^{m} e_{ji}a_{ik} = a_{ik},$$
$k$ ranging from 1 to $n$. In the other rows all entries are zero. Hence $E_{ji}A = A_{ij}$, which has the $j$th row corresponding to the $i$th row of $A$, and zeros elsewhere. By similar calculations we obtain $E_{jj}A = A_{jj}$ and $E_{ij}A = A_{ji}$. Hence
\[ P_{ij}A = A - A_{ii} + A_{ij} - A_{jj} + A_{ji}. \]

Thus \( P_{ij} \) produces operation (i) on \( A \).

(ii) \( (M_i(c))A = (I - E_{ii} + cE_{ii})A 
\]
\[ = A - A_{ii} + cA_{ii}, \]
by similar calculations. Thus \( (M_i(c))A \) produces operation (ii) on \( A \).

(iii) \( A_{ij}A = (I + E_{ji})A = A + A_{ij}. \)
Thus \( A_{ij}A \) produces operation (iii) on \( A \). This completes the proof.

Recall that by definition, an \( n \times n \) matrix \( A \) is non-singular if and only if a matrix \( B \) exists such that \( AB = I \).
Moreover, by Theorem 4.5 (p. 78, Finkbeiner) \( AB = I \) implies \( BA = I \). Hence, setting \( B = A^{-1} \), we have:

\[ A \text{ is non-singular if and only if } AA^{-1} = A^{-1}A = I. \]

Recall further that by Theorem 4.6 (p. 78, Finkbeiner) the following statements are equivalent for an \( n \times n \) matrix \( A \):

(i) \( A \) is non-singular;
(ii) the row vectors of \( A \) are linearly independent;
(iii) \( A \) has rank \( n \).

**Lemma 2.** The identity matrix \( I \) is non-singular.

**Proof.** Let \( X_I \) denote the linear combination of the \( m \) rows of \( I \), with scalars \( x_k^{'} \), \( k = 1, \ldots, m \). Then

\[ X_I = x_1(1, \ldots, 0) + x_2(0, 1, 0, \ldots, 0) + \ldots + x_m(0, \ldots, 0, 1) \]
\[ = (x_1, \ldots, 0) + (0, x_2, 0, \ldots, 0) + \ldots + (0, \ldots, 0, x_m) \]
= \{x_1, x_2, \ldots, x_m\}
= \{0, 0, \ldots, 0\} \text{ if and only if } x_k = 0 \text{ for all } k.

Hence the m rows of I are linearly independent. Thus I is non-singular, by Theorem 4.6 above.

**Theorem 2.** Every elementary matrix is non-singular.

**Proof.** We use the notation of Lemma 2. Consider $P_{ij}$. Permuting the $i$th and $j$th rows of I does not affect the linear independence of its rows; thus

$X_{p_{ij}} = (x_1, \ldots, x_j, \ldots, x_i, \ldots, x_m) = (0, \ldots, 0)$ if and only if each $x_k = 0$.

Consider $M_{ij}(c)$. Replacing the $(i,i)$ entry of I by a non-zero scalar $c$,

$X_{M_{ij}}(c) = (x_1, \ldots, cx_i, \ldots, x_m) = (0, \ldots, 0)$ if and only if $x_k = 0$ for all $k$.

Consider $A_{ij}$. Replacing $x_i$ by $x_i + x_j$, we have

$X_{A_{ij}} = (x_1, \ldots, x_i + x_j, \ldots, x_j, \ldots, x_m) = (0, \ldots, 0)$ if and only if $x_1 = \ldots = x_i + x_j = \ldots = x_j = \ldots = x_m = 0$.

But $x_j = 0$ implies $x_i = 0$.

Hence the m rows of an elementary matrix of each type are linearly independent, so that each is non-singular.

The important consequence of this theorem is that every elementary matrix has an inverse.
Lemma 3. $E_{ik}E_{kj} = \delta_{kh}E_{ij}$. Therefore each $E_{ij}$ is either idempotent or nilpotent of index 2, according to whether $i=j$ or $i \neq j$.

Proof. Recall the meaning of $E_{rs}$. Also, $\delta_{kh} = \begin{cases} 1 & \text{if } k = h \\ 0 & \text{if } k \neq h \end{cases}$

Suppose $k = h$. The entry in the $(i,j)$ position of $E_{ih}E_{hj}$ is $\sum_{h=1}^{m} e_{ih}e_{hj} = e_{ij}$ since all the terms in this sum are zero except one. As $e_{ij}$ is the only non-zero entry, the other rows are zero. Hence

$$E_{ih}E_{hj} = E_{ij} = \delta_{hh}E_{ij}.$$  

If $k \neq h$, $e_{ik}$ and $e_{nj}$ are not factors of a product in any row or column of $E_{ik}E_{kj}$; hence $E_{ik}E_{kj} = Z$, the zero matrix. Also $\delta_{kh}E_{ij} = 0(E_{ij}) = Z$.

Now $E_{ij}E_{jj} = \delta_{jj}E_{jj} = E_{jj}$ so that $E_{ij}$ is idempotent for $i = j$. If $i \neq j$, $E_{ij}E_{ij} = \delta_{ij}E_{ij} = 0(E_{ij}) = Z$.

Since $E_{ij} \neq Z$, $E_{ij}$ is nilpotent of index 2. This proves the lemma.

Theorem 3. The inverse of an elementary matrix of type P or type M is an elementary matrix of the same type. The inverse of an elementary matrix of Type A is the product of elementary matrices of type M and type A.

Proof. Consider the product $P_{ij}P_{ij}$. Using Lemma 3,

$$P_{ij}P_{ij} = (I - E_{ii} + E_{ij} - E_{jj} + E_{ij})(I - E_{ii} + E_{ij} - E_{jj} + E_{ij})$$

$$= I - E_{ii} + E_{ii} - E_{jj} + E_{ij} - E_{ii} + E_{ii} - E_{ii} + E_{ii} - E_{ij}$$

$$- E_{ij} + E_{ij} - E_{ij} + E_{ij} + E_{ij} + E_{ij} + E_{ij} - E_{ij}$$
Consider $M_i(c^{-1})$.

$$(M_i(c))(M_i(c^{-1})) = (I + cE_{ii} - E_{ii})(I + c^{-1}E_{ii} - E_{ii})$$

$$= I + c^{-1}E_{ii} - E_{ii} + cE_{ii} + E_{ii} - cE_{ii}$$
$$- E_{ii} - c^{-1}E_{ii} + E_{ii}$$

$$= I,$$ so that $M_i(c^{-1})$ is the inverse of $M_i(c)$.

We prove that $I - E_{ji}$ is the inverse of $I + E_{ji} = A_{ij}$:

$$(I + E_{ji})(I - E_{ji}) = I - E_{ji} + E_{ji} = I.$$ We now prove the last part of the theorem; namely, that

$$I - E_{ji} = (M_i(-1)) A_{ij}(M_i(-1)):$$

$${M_i(-1)}A_{ij}{M_i(-1)} = (I - E_{ii} - E_{ii})(I + E_{ji})(I - E_{ii} - E_{ii})$$
$$= (I - 2E_{ii})(I + E_{ji})(I - 2E_{ii})$$
$$= (I + E_{ji} - 2E_{ii})(I - 2E_{ii})$$
$$= I - 2E_{ii} + E_{ji} - 2E_{ji} + 2E_{ii}$$
$$= I - E_{ji}.$$
II. ROW EQUIVALENCE

Elementary row operations offer a convenient method of comparing two matrices.

**Definition.** An \( m \times n \) matrix \( B \) is said to be **row equivalent** to an \( m \times n \) matrix \( A \) if and only if \( B \) can be obtained by performing a finite number of elementary row operations on \( A \).

**Lemma 4.** If \( B \) is row equivalent to \( A \), then \( B = E_k \cdots E_2 E_1 A \) for suitable elementary matrices \( E_i \), \( i = 1, 2, \ldots, k \).

**Proof.** By Theorem 1, \( E_1 A \) represents an elementary row operation on \( A \). \( E_2 E_1 A \) represents a second such operation. Hence \( E_k \cdots E_2 E_1 A \) represents a succession of \( k \) elementary row operations on \( A \). By definition, \( B \) is row equivalent to \( A \).

**Lemma 5.** Row equivalence of matrices is an equivalence relation.

**Proof.** Let \( A, B, \) and \( C \) be \( m \times n \) matrices. Define \( A \sim B \) to mean that \( A \) is row equivalent to \( B \).

(i) For scalar multiple 1, \( M_1(c) = M_1 (I) = I \). Hence \( A = IA = M_1(I)A \) so that \( A \sim A \). Therefore \( \sim \) is reflexive.

(ii) If \( A \sim B \), then by Lemma 4, \( A = E_p \cdots E_{p-k} B \), for suitable \( E_i \). By Theorem 2, an elementary matrix is
non-singular hence has an inverse. By a succession of inverse operations, we obtain \( B = E_{p-k}^{-1} \ldots E_{1}^{-1} A \). By Theorem 3, \( E_{p-k}^{-1} \ldots E_{1}^{-1} \) is a product of elementary matrices, hence \( B \sim R \sim A \). Thus \( R \) is symmetric.

(iii) If \( A \sim R \sim B \) and \( B \sim R \sim C \), we have \( A = E_{p} \ldots E_{p-k}B \) and \( B = E_{k} \ldots E_{1}C \), for suitable \( E_{i}, i=1, \ldots, p \). Then \( A = E_{p} \ldots E_{p-k}(E_{k} \ldots E_{1}C) = E_{p} \ldots E_{p-k}E_{k} \ldots E_{1}C \), since matrix multiplication is associative. Thus \( A \sim R \sim C \) so that \( R \) is transitive. Hence \( R \) is an equivalence relation.

Given \( A \) row equivalent to \( B \), we can now say that \( A \) and \( B \) (or \( B \) and \( A \)) are row equivalent.

**Theorem 4.** Row equivalent matrices have the same rank.

**Proof.** Let \( r(P) \) denote the rank of matrix \( P \). Recall Theorem 4.7 (p. 78, Finkbeiner) by which if matrix \( P \) is non-singular, then for any matrix \( Q \), \( r(PQ) = r(Q) = r(QP) \).

If \( B \) is row equivalent to \( A \), then \( B = E_{k} \ldots E_{1}A \), for suitable \( E_{i} \). By Theorem 4.8 (p. 78, Finkbeiner), the product of non-singular matrices is non-singular, hence \( E_{k} \ldots E_{1} \) is a non-singular matrix. Hence \( r(B) = r(E_{k} \ldots E_{1}A) = r(A) \), as was to be proved.

**Definition.** An \( m \times n \) matrix is said to be in **echelon form** if it satisfies the following conditions:
(i) The first $k$ rows are non-zero; the other rows are zero, and

(ii) The first non-zero element in each non-zero row is 1, and it appears in a column to the right of the first non-zero element of any preceding row.

**Definition.** Any matrix in echelon form is said to be in **reduced echelon form** (r.e.f.) if it has the additional property:

(iii) The first non-zero element in each non-zero row is the only non-zero element in its column.

**Theorem 5.** Any $m \times n$ matrix of rank $k$ is row equivalent to a matrix in echelon form (also, reduced echelon form) with $k$ non-zero rows.

**Proof.** Let $A = (a_{ij})$ be an $m \times n$ matrix of rank $k$. We perform the following elementary row operations on $A$:

- Suppose the $q$th column is the first column in which a non-zero element, say $a_{pq}$, appears. If we multiply row $p$ by $a_{pq}$, then interchange row $p$ and row 1, we obtain 1 in the $(1,q)$ position. By adding suitable multiples of the first row to the other rows, we obtain zeros in column $q$ below the first row. Denote the resulting matrix by $B$. $B$ has the form
where each * denotes some scalar and where R is an (m-l) x (n-q) submatrix.

We continue the process of operating with the last m-l rows of B: in R we find the first column in which a non-zero element \( b_{rs} \) appears, and multiply row r by \( b_{rs}^{-1} \), then interchange row r with row 2 of B, to obtain 1 in the (2,s) position. By adding suitable multiples of the second row to the other rows, we produce zeros in column s in every row below row 2. We repeat these operations until we obtain a matrix C in echelon form.

Now consider the non-zero rows of C. Suppose that \( c_{2j}=1 \) is the first non-zero entry in row 2. If \( c_{1j} \neq 0 \), we add to row 1 a suitable multiple of row 2 to obtain \( c_{1j}=0 \). Suppose that \( c_{3k}=1 \) is the first non-zero element in row 3. Then we add suitable multiples of row 3 to the first two rows to produce \( c_{1k}=c_{2k}=0 \). We repeat the process in every column containing non-zero elements above the first non-zero entry of a given row. Eventually all entries above the leading 1's are replaced by zeros. The resulting matrix, say D, is now in reduced echelon form, and row equivalent to A.
We still need to prove that $D$ has $K$ non-zero rows. Since $D$ is row equivalent to $A$, by Theorem 4 $r(D) = r(A) = k$. Since the rank of a matrix is the maximal number of linearly independent rows, $k+1$ non-zero rows would be linearly dependent. If the non-zero rows numbered less than $k$, $r(D) < k$. Hence $D$ has exactly $k$ non-zero rows. This proves the theorem.

**Lemma 6.** The transpose of any elementary matrix is an elementary matrix.

**Proof.** Recall that by Theorem 4.1 (p. 71, Pinkbeiner),

(i) $(A+B)' = A' + B'$ and (ii) $(AB)' = B'A'$, where if $A = (a_{ij})$ the transpose of $A$, denoted $A' = (a_{ji})$. (We need part (ii) above to prove the next theorem). Then

$$(P_{ij})' = (I - E_{ii} + E_{ji} - E_{jj} + E_{ij})'$$

$$= I - E_{ii} + E_{ji} - E_{jj} + E_{ij} = P_{ij}.$$

$$(M_{i}(c))' = (I - E_{ii} + cE_{ii})' = M_{i}(c).$$

$$(A_{ij})' = (I + E_{ij})' = I + E_{ij} = A_{ji}.$$

**Theorem 6.** The number of linearly independent columns of any matrix $A$ equals the number of linearly independent rows of $A$. Thus $r(A) = r(A')$.

**Proof.** Let $E$ be a reduced echelon matrix and row equivalent to $A$. By Theorem 4, if $r(A) = k$, $R(E) = k$. By Theorem 5, $E$ has $k$ non-zero rows each of which has a 1
as the first non-zero element. By definition, each 1 is in a column of E where it is the only non-zero element. These k columns are linearly independent, and a (k+1)st column is a linear combination of these. The columns of E are the rows of $E'$. Hence $r(E') = k$. For suitable $E_1$, $E = E_k \ldots E_1 A$. Taking a transpose of both sides of the equation, $E^t = A^t E_1^t \ldots E_k^t$. By Lemma 6, $E_1^t \ldots E_k^t$ is a product of elementary matrices, hence non-singular. Thus $r(E') = r(A'E_1^t \ldots E_k^t) = r(A') = k$.

**Theorem 7.** A square matrix is non-singular if and only if it is row equivalent to the identity matrix.

**Proof.** An $n \times n$ identity matrix $I$ has rank $n$ since it is non-singular, by Lemma 2. Let $A$ be an $n \times n$ matrix and be row equivalent to $I$. Then $A$ has rank $n$, hence is non-singular.

Conversely, if $A$ is non-singular it has rank $n$. Let $A$ be row equivalent to some $n \times n$ matrix $E$ in reduced echelon form. Then $E$ has rank $n$. Hence $E$ must contain $n$ non-zero rows such that the first non-zero element of each row is 1. Thus there are $n$ columns each with 1 as the only non-zero element. These are the properties of the identity matrix; hence $E = I$.

**Theorem 8.** A square matrix is non-singular if and only if it is the product of elementary matrices.
Proof. If $A$ is non-singular, then by Theorem 7 and Lemma 5, $A = E_k \ldots E_1 I = E_k \ldots E_1$, for suitable elementary matrices $E_i$.

Conversely, let $A = E_k \ldots E_1$, for suitable $E_i$. Since by Theorem 2 an elementary matrix is non-singular, their product is non-singular, by Theorem 4.8 (p. 78, Finkbeiner). Hence $A$ is non-singular.

If $A$ is non-singular, $A$ and $I$ are row equivalent so that for suitable $E_i$, $E_p \ldots E_1 A = I$. Multiplying each side of the equation on the right by $A^{-1}$, we obtain $E_p \ldots E_1 I = A^{-1}$. Hence the same sequence of elementary row operations which reduces $A$ to $I$, changes $I$ to $A^{-1}$. If $A$ is singular, these operations will reduce $A$ to a singular row equivalent matrix.

Theorem 9. Matrix $B$ is row equivalent to matrix $A$ if and only if $B = PA$ for some non-singular matrix $P$.

Proof. If $B$ is row equivalent to $A$, $B = E_k \ldots E_1 A$ for suitable $E_i$. Since each $E_i$ is non-singular, their product is a non-singular matrix. Hence by Theorem 8 there exists a non-singular matrix $P$ such that $P = E_k \ldots E_1$. Then $B = PA$.

Conversely, let $B = PA$ where $P$ is non-singular. By Theorem 8, $P$ is a product of elementary matrices, say $E_k \ldots E_1$. Then $B = E_k \ldots E_1 A$. Hence $B$ is row equivalent
Lemma 7. If B is row equivalent to A, then any row of B is a linear combination of the rows of A.

Proof. Let \( B = E_k \ldots E_1 A \), for suitable elementary matrices \( E_i \). Recall that premultiplication of A by each \( E_i \) performs an elementary row operation on A. For example, if \( E_2 E_1 A \) adds row i to row j, then multiplies the row by a non-zero scalar c, then the \( j^{th} \) row of \( E_2 E_1 A \) is \( cA_i + cA_j \). After a succession of \( k \) elementary row operations on A, the \( i^{th} \) row of \( E_k \ldots E_1 A \) has the form \( c_1 A_i + c_2 A_2 + \ldots + c_mA_m \), for some scalars \( c \). But this equals the \( i^{th} \) row of B. This proves the lemma.

Theorem 10. Relative to a pair of bases, let matrices A and B represent linear transformations \( T \) and \( S \). Then A and B are row equivalent if and only if

\( R_T = R_S \) (the range spaces of \( T \) and \( S \)).

Proof. Let \( A = (a_{ij})_{mxn} \), and \( B = (b_{ij})_{mxn} \) determine the linear transformations \( T \) and \( S \) from \( V_m \) with \( \alpha \)-basis to \( W_n \) with \( \beta \)-basis. Then \( \alpha_i T \) and \( \alpha_i S \) are respectively represented in the \( \beta \)-basis by the \( i^{th} \) rows of A and B, \( i=1,\ldots,m \). Then \( R_S \) is the subspace of \( W_n \) which is spanned by the rows of B.

Let A be row equivalent to B. Then by Lemma 7, each row of A is a linear combination of the rows of B.
Therefore each vector $\alpha_i T$ is a linear combination of the vectors $\alpha_i S$. Hence $R_T$ is a subspace of $R_S$. By Lemma 5, $B$ is row equivalent to $A$. Thus, similarly, each vector $\alpha_i S$ is a linear combination of the vectors $\alpha_i T$. Hence $R_S$ is a subspace of $R_T$. Hence $R_T = R_S$.

Conversely, suppose $R_T = R_S$. Then each $\alpha_i T$ is a linear combination of the $\alpha_i S$. This implies that each row of $A$ is a linear combination of the rows of $B$. Since linear combination of rows can be performed by elementary row operations, $A$ is row equivalent to $B$. By Lemma 5, $B$ is row equivalent to $A$, so that $A$ and $B$ are row equivalent. This completes the proof.

**Theorem 11.** Two matrices $A$ and $B$ in reduced echelon form are row equivalent if and only if $A = B$.

**Proof.** If $A = B$, then by the reflexive property of Lemma 5, $A$ and $B$ are row equivalent.

Conversely, let $A$ and $B$ be in reduced echelon form (r.e.f.), and row equivalent. Then $r(A) = r(B) = k$, say, so that each matrix has $k$ linearly independent rows and the first $k$ rows of each matrix are non-zero. Hence for $i \leq k$, the $i^{th}$ row of $B$ is a linear combination of the $k$ non-zero rows of $A$:

$$B_i = \sum_{j=1}^{k} c_j A_j$$

where not all $c_j$ are zero since $B_i$ is a non-zero row.
Since A and B are in r.e.f., the first non-zero entry of row i is 1, and it appears in, say, column $t_i$ of A and column $s_i$ of B, for every $i \leq k$. By definition, in columns $s_i$ and $t_i$ all other entries are zero. From the above equation, the $s_i$-entry of $B_i$ is a linear combination of the entries in column $s_i$ of matrix A. Since the $s_i$-entry of $B_i$ is 1, for some $j \neq k$ there is at least one non-zero entry in the $s_i$-column of A, with the corresponding $c_j \neq 0$.

We wish to determine the position of $t_j$ relative to $s_i$. The first non-zero entry of $A_j$ appears in the $t_j$ position. If $t_j > s_i$, the $t_j$-entry cannot be the first non-zero entry since it is preceded by the non-zero $s_i$-entry. Hence $t_j \leq s_i$. Since A and B are row equivalent, we can reverse the roles of A and B. Hence, by similar reasoning, if $t_j < s_i$, then the $s_i$-entry is not the first non-zero element. Therefore $s_i = t_j$ for some $j$. Hence the ordered sets of integers, $s_1 < s_2 < \ldots < s_k$ and $t_1 < t_2 < \ldots < t_k$, are the same set, and the ordering implies that $s_i = t_i$, $i = 1, \ldots, k$. Hence the first non-zero entries in the rows of A and B are in identical positions.

In row $B_i$, any entry before the $i^{th}$ position is zero, so that the linear combination of the components of the corresponding column of A is zero. Hence $c_j = 0$.
for \( j=1,2,\ldots,i-1 \). In the \( i^{\text{th}} \) position of \( B_i \), the entry is either 1 or zero, so that the corresponding \( c_i = 1 \).

Hence

\[
B_i = A_i + \sum_{j=i+1}^{k} c_j A_j.
\]

In terms of the elements of \( A \) and \( B \), we have

\[
b_{ir} = a_{ir} + \sum_{j=i+1}^{k} c_j a_{jr}.
\]

Next we show that \( c_j a_{jr} = 0 \) for all \( i+1 \leq j \leq k \).

Suppose \( a_{qr} \neq 0 \) for some \( q \), \( i < q \leq k \). Since the first non-zero entry in row \( q \) of \( A \) occurs in the \( t_q \) position,

\[
a_{qt_q} = 1 \quad \text{for some} \quad t_q, \quad t_i < t_q \leq r.
\]

The non-zero entries of \( A \) and \( B \) are in identical positions, hence \( b_{qt_q} = 1 \).

But other entries in column \( t_q \) are zero, so that

\[
b_{pt_q} = 0 = a_{pt_q} \quad \text{if} \quad p \neq q. \quad \text{In particular,} \quad i \neq q. \quad \text{Hence}
\]

\[
b_{it_q} = a_{it_q} + \sum_{j=i+1}^{k} c_j a_{jt_q} = 0 + c_q a_{qt_q}.
\]

Since \( a_{qt_q} = 1 \), \( c_q = 0 \). Therefore for \( i+1 \leq q \leq k \), if \( a_{qr} \neq 0 \), then \( c_q = 0 \). Hence \( b_{ir} = a_{ir} \) for all \( r \) and for \( i \leq k \). For \( i > k \), all the entries are zero by definition.

Hence all the corresponding entries of \( A \) and \( B \) are equal and \( A = B \).

This completes the proof.

The result of Theorem 11 is that there is exactly one matrix in reduced echelon form which is row
equivalent to any given matrix; that is, reduced echelon form is canonical with respect to row equivalence.

We noted in the preceding section that a system of linear equations is reducible to an equivalent system by elementary operations. Since these operations correspond to the elementary row operations on a matrix, row equivalent augmented matrices represent equivalent systems of linear equations. Specifically, recall the system of equations given by (2),

$$\sum_{j=1}^{n} a_{ij} x_j = c_i, \ i = 1, \ldots, m.$$  

In matrix form, we have $A X = C$ where $A = (a_{ij})_{mxn}$, and the column vectors $X = (x_j)_{nx1}$ and $C = (c_i)_{mx1}$. If $C$ is adjoined to $A$, we get the augmented matrix $(A, C)_{mx(n+1)}$. Operations on the augmented matrix correspond to the operations changing the system (2) to an equivalent system; hence $A X = C$ and $A^* X = C^*$ have the same solutions if their augmented matrices $(A, C)$ and $(A^*, C^*)$ are row equivalent.

If in the above each $c_i = 0$, we get a homogeneous system of equations, in which case $A X = 0$ and $A^* X = 0$ have the same solutions if $A$ and $A^*$ are row equivalent.
III. EQUIVALENCE

Elementary operations may also be performed on columns. If we replace the word "row" by the word "column," we obtain a definition for the three elementary column operations.

**Lemma 8.** An elementary column operation on a matrix $A$ is the same as an elementary row operation on $A'$; the column operation may be performed by post-multiplying $A$ by an elementary matrix.

**Proof.** Let $EA' = B$, where $E$ is an $m \times m$ elementary matrix and $A'$ is the transpose of an $n \times m$ matrix $A$. Taking a transpose of both sides of the equation,

$$B' = (EA') = (A')'E' = A E'.$$

Recall that the $i^{th}$ row of $B$ is the $i^{th}$ column of $B'$. Hence if $EA'$ is the result of an elementary row operation on the $i^{th}$ row of $A'$, $AE'$ is the result of an elementary column operation on the $i^{th}$ column of $A$. By Lemma 6, $E'$ is an elementary matrix. Since $AE'$ denotes post-multiplying $A$ (i.e., multiplying on the right) by $E'$, the lemma is proved.

The results established for row equivalence in the preceding section are valid for column equivalence. They can be deduced by a substitution of Lemma 8 for Theorem 1 in the proofs and a replacement of the word
"row" by the word "column" throughout the discussion.

We now consider the effect on a matrix when both elementary row and column operations are performed on A.

**Definition.** Matrix B is said to be equivalent to matrix A if and only if B can be obtained by performing a finite number of elementary row and column operations on A.

**Theorem 12.** B is equivalent to A if and only if $B = PAQ$ for suitable non-singular matrices P and Q.

**Proof.** Let B be any $m \times n$ matrix. If B is equivalent to A, then by definition there exist $m \times m$ elementary matrices $E_i$ and $n \times n$ elementary matrices $F_j$ such that

$$B = E_k \ldots E_1 A F_1 \ldots F_p, \quad i = 1, \ldots, k; \quad j = 1, \ldots, p.$$  

Since $E_i$ and $F_j$ are non-singular, their respective products are non-singular. Let $E_k \ldots E_1 = P$ and $F_1 \ldots F_p = Q$. Then $B = PAQ$, where P and Q are non-singular matrices.

Conversely, suppose $B = PAQ$ for suitable non-singular matrices P and Q. By Theorem 8, P and Q are products of elementary matrices, say the above $E_i$ and $F_j$. Then $B = E_k \ldots E_1 A F_1 \ldots F_p$. Hence, by definition, B is equivalent to A.

**Theorem 13.** Equivalence of matrices is an equivalence relation.
Proof. Let A, B, and C be m x n matrices. Let A R B mean that A is equivalent to B. By Theorem 12, A R B if and only if A = PBQ, where P and Q are respective m x m and n x n non-singular matrices.

(i) If P = I and Q = I are, respectively, m x m and n x n identity matrices, then A = IAI = PAQ. Hence A R A so that R is reflexive.

(ii) If A R B, A = PBQ. Let P^{-1} and Q^{-1} be inverses of the non-singular matrices P and Q. Multiplying the equation by P^{-1} on the left and by Q^{-1} on the right, we obtain B = P^{-1}AQ^{-1}. Since P^{-1} and Q^{-1} are non-singular matrices, B R A. Hence R is symmetric.

(iii) If A R B and B R C, A = PBQ and B = RCS for non-singular m x m matrices P and R and n x n matrices Q and S. Then A = P(RCS)Q = (PR)C(SQ) where the products PR and SQ are non-singular m x m and n x n matrices. Therefore A R C, so that R is transitive. Hence R is an equivalence relation.

If A is equivalent to B, we can now say A and B (or B and A) are equivalent.

Theorem 14. An m x n matrix of rank k is equivalent to the m x n matrix B in which b_{11} = b_{22} = \ldots = b_{kk} = 1, and b_{ij} = 0 otherwise.

Proof. Let A be an m x n matrix of rank k. We perform
the elementary row and column operations indicated. If $k = 0$, $A = Z$ and there is nothing to prove. If $k \neq 0$, we move some non-zero $a_{ij}$ to position $(1,1)$ by interchanging rows $i$ and $1$ and columns $j$ and $1$. Then we replace $a_{ij}$ by $1$ by multiplying the first row by $a_{ij}^{-1}$. Next we obtain zeros in the remainder of row $1$ by adding to each of the $n-1$ columns a suitable multiple of the first column. Similarly, we obtain zeros in the remainder of column $1$ by adding suitable multiples of the first row to each of the $m-1$ rows. We now have a matrix $C$, say, in the form

$$
C = PAQ = \begin{pmatrix}
1 & 0 & \ldots & 0 \\
0 & \ddots & R \\
\vdots & \ddots & \ddots \\
0 & \ldots & 1 & 0
\end{pmatrix}
$$

where $P$ is the product of elementary matrices corresponding to the row operations, and $Q$ is the product of elementary matrices corresponding to the column operations. Hence $C$ is equivalent to $A$.

If the $(m-1) \times (n-1)$ submatrix $R = Z$, the theorem is proved. Otherwise we fix the first row and column of $C$ and repeat similar operations on the $(m-1)$ rows and the $(n-1)$ columns of $R$. That is, we multiply the $k^{th}$ row containing a non-zero $c_{hk}$ by $c_{hk}^{-1}$ then interchange rows and columns to obtain the $1$ in the $(2,2)$ position of $C$. By adding suitable multiples of row $2$ to the $m-2$
rows and suitable multiples of column 2 to the n-2 columns, we produce zeros in the remainder of the second row and column.

If the remaining \((m-2) \times (n-2)\) submatrix \(S \neq Z\), we continue until, through a finite number of row and column operations, the remaining submatrix \(X = Z\). The resulting matrix \(B\) is such that \(A = EBF\), where \(E\) and \(F\) are the respective products of \(m \times m\) and \(n \times n\) elementary matrices, which represent the elementary row and column operations; hence \(E\) and \(F\) are non-singular matrices. Thus \(A\) is equivalent to \(B\).

We reduced \(B\) to having all zero entries except for a certain number of 1's in the \((i,i)\) position. We now prove that \(i = 1, \ldots, k = r(A)\): the number of linearly independent rows of \(B\) is precisely the number of rows having 1 in the \((i,i)\) position (i.e., the number of non-zero rows). This number is the rank of \(B\). By Theorem 4.7 (p. 78, Finkbeiner), \(r(A) = r(EBF) = r(B) = k\). This completes the proof.

The result of Theorem 14 is that every \(m \times n\) matrix of rank \(r\) is reducible to the same canonical form 
\[
\begin{pmatrix}
I & Z \\
Z & Z
\end{pmatrix}
\]
where \(I\) is the \(r \times r\) identity submatrix, and \(Z\) are the zero submatrices of suitable dimensions. Hence the form \(\begin{pmatrix}I & Z \\ Z & Z\end{pmatrix}\) is canonical with respect to equivalence.
of matrices.

Theorem 15. Two \( m \times n \) matrices are equivalent if and only if they have the same rank.

Proof. Let \( A \) and \( C \) be \( m \times n \) matrices. If \( A \) and \( C \) are equivalent, then \( A = PCQ \) where \( P \) and \( Q \) are non-singular matrices. Hence \( r(A) = r(PCQ) = r(C) \).

Conversely, if \( r(A) = r(C) \) each is equivalent to matrix \( B \) of Theorem 14. Hence by the transitive property of Theorem 13, \( A \) and \( B \) are equivalent.

Theorem 16. A square matrix is non-singular if and only if it is equivalent to the identity matrix.

Proof. Let \( A \) be an \( m \times m \) matrix. If \( A \) is equivalent to the \( m \times m \) matrix \( I \), by Theorem 15 \( r(I) = r(A) = m \). Then \( A \) has \( m \) linearly independent rows, hence is non-singular.

Conversely, if \( A \) is non-singular, it has rank \( m \). Since for an \( m \times m \) identity matrix \( I \), \( r(I) = m \), by Theorem 15 matrix \( A \) is equivalent to the identity matrix.

As a result of Theorem 16, a non-singular matrix \( A \) may be reduced to the identity matrix \( I \) by a succession of elementary row and elementary column operations.

Since \( I \) and \( A \) are equivalent, \( I = E_k \ldots E_1 A F_1 \ldots F_p \).

Denote \( I = PAQ \). Then, multiplying by \( P^{-1} \) and \( Q^{-1} \), \( P^{-1}Q^{-1} = A \) and taking the inverse of the equation,
$A^{-1} = QP = QIP$. This suggests still another way of calculating $A^{-1}$.\[\begin{array}{c|c}
I & A \\
(\text{a}) & \\
(\text{b}) & I
\end{array}\]
Set up an array \[\begin{array}{c|c}
I & A \\
(\text{a}) & \\
(\text{b}) & I
\end{array}\]. Reduce $A$ to $I$ by elementary row and column operations as follows: perform the same row operation on $I$ in quadrant (a) as on $A$; perform the same column operation on $I$ in quadrant (b) as on $A$. A finite sequence of these operations will change the above array to \[\begin{array}{c|c}
Q & I \\
(\text{a}) & \\
(\text{b}) & P
\end{array}\]. Then $A^{-1} = QP$.

We note that the canonical form, with respect to equivalence, for an $n \times n$ non-singular matrix $A$ is a special case of $\begin{pmatrix} I & Z \\ Z & Z \end{pmatrix}$, since in this case the number of linearly independent rows (hence columns) is $n$. Hence for $A$, $\begin{pmatrix} I & Z \\ Z & Z \end{pmatrix} = I$, the $n \times n$ identity matrix.
IV. SIMILARITY

In Section I, we derived a matrix representation of a linear transformation $T$ by fixing the basis vectors of the domain $V_m$ and the range $W_n$ of $T$. With this stipulation, $T$ was represented uniquely by a matrix, but this matrix representation depended on the choice of bases.

We now wish to determine the relationship between two matrices each of which represents the same linear transformation $T$ relative to an independent choice of bases for $V_m$ and $W_n$. This is given by the next theorem.

**Theorem 17.** Two $m \times n$ matrices $A$ and $C$ represent the same linear transformation from $V_m$ to $W_n$ relative to two suitably chosen pairs of bases if and only if $A$ and $C$ are equivalent.

The proof, as given by Finkbeiner (pp. 121-122), consists of choosing two pairs of bases for $V_m$ and $W_n$, as follows: with respect to $\{\alpha\}$ and $\{\beta\}$ bases, $T$ is represented by a uniquely determined matrix $A$. With respect to $\{\gamma\}$ and $\{\delta\}$ bases, $T$ is represented by a uniquely determined matrix $C$. Then

$$\alpha_i T = \sum_{j=1}^{n} a_{ij} \beta_j \quad \text{and} \quad \gamma_j T = \sum_{k=1}^{n} c_{jk} \delta_k. $$

Let $R$ map $\gamma_i$ onto $\alpha_i$ in $V_m$. By Theorem 3.6, (p. 55), $R$ is non-singular since it maps a basis onto a basis,
and relative to the $\gamma_i$ basis it is represented uniquely by a non-singular matrix $P$, where

$$\alpha_i = \gamma_i = \sum_{j=1}^{m} p_{ij} \gamma_j.$$  

Similarly, let $S$ map $\beta_j$ onto $\delta_j$ in $W_n$. Then $S$ is represented by a uniquely determined non-singular matrix $Q$, where

$$\beta_j = \delta_j = \sum_{k=1}^{n} q_{jk} \delta_k.$$  

Calculating $\alpha T$ in two different ways, we obtain

$$\alpha_i T = \sum_{j=1}^{n} a_{ij} \beta_j = \sum_{j=1}^{n} a_{ij} \left( \sum_{k=1}^{n} q_{jk} \delta_k \right)$$

$$= \sum_{k=1}^{n} \left( \sum_{j=1}^{n} a_{ij} q_{jk} \right) \delta_k.$$  

$$\alpha_i T = \left( \sum_{j=1}^{m} p_{ij} \gamma_j \right) T = \sum_{j=1}^{m} p_{ij} \left( \sum_{k=1}^{n} \delta_k \right)$$

$$= \sum_{k=1}^{n} \left( \sum_{j=1}^{m} p_{ij} \delta_k \right) \delta_k, \text{ for } i = 1, \ldots, m.$$  

Since $\alpha_i T$ is a unique linear combination of $\delta_k$,

$$\sum_{j=1}^{n} a_{ij} q_{jk} = \sum_{j=1}^{m} p_{ij} c_{jk}.$$  

This equation represents the entries in the $(i,k)$ position of $AQ$ and $PC$, so that $AQ = PC$, or

$$A = PCQ^{-1}$$

where $P$ and $Q^{-1}$ are $m \times m$ and $n \times n$ non-singular matrices. Hence $A$ and $C$ are equivalent.

Conversely, let $A$ and $C$ be equivalent matrices. Then, without loss of generality, $A = PCQ^{-1}$, for suitable non-singular matrices $P$ and $Q^{-1}$. Hence $AQ = PC$. Choose
any basis pair \{\gamma\} for \(V_m\) and \{\delta\} for \(W_n\) and let \(T\) be the linear transformation represented by \(C\) relative to this choice of bases. Let \(R\) and \(S\) be the linear transformations as before, represented by the non-singular matrices \(P\) and \(Q\). Then \(T \alpha_{ij}^\gamma \beta_{ij}^\delta = R \gamma_{ij} T \alpha_{ij}^\gamma \beta_{ij}^\delta \).

\[
\gamma_{iTS} = \left(\sum_{j=1}^{n} c_{ij} \delta_{j}\right) S = \sum_{j=1}^{n} c_{ij} \left(\sum_{k=1}^{n} q_{jk} \delta_{k}\right) = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} c_{ij} q_{jk}\right) \delta_{k} = \alpha_{i} T \text{ by previous calculation.}
\]

Also,

\[
\gamma_{iRT} = \left(\sum_{j=1}^{m} p_{ij} \gamma_{j}\right) T = \sum_{j=1}^{m} p_{ij} \left(\sum_{k=1}^{n} c_{jk} \delta_{k}\right) = \sum_{k=1}^{n} \left(\sum_{j=1}^{m} p_{ij} c_{jk}\right) \delta_{k} = \alpha_{i} T \text{ by previous calculation.}
\]

Hence \(A\) and \(C\) represent the same linear transformation relative to the different pairs of bases. This completes the proof.

If the roles of matrices and linear transformations in the above theorem are reversed, we have an analogous theorem for linear transformations:

**Theorem 18.** Two linear transformations, \(T_1\) and \(T_2\), are represented relative to two suitably chosen pairs of bases by the same matrix if and only if non-singular linear transformations \(R\) and \(S\) exist such that \(T_2 = R T_1 S^{-1}\).

To prove the theorem, consider a single \(m \times n\) matrix \(A\) and two pairs of bases, \(\alpha, \beta\) and \(\gamma, \delta\). Relative to each basis pair, \(A\) determines a linear transformation,
say $T_1$ and $T_2$. Define $R$ and $S$ as in Theorem 17; that is, 
$\gamma_i R = \alpha_i$ and $\delta_j S = \beta_j$. Then

$$
\gamma_i T_2 = \sum_{j=1}^{n} a_{ij} \delta_j \quad \text{and} \quad \alpha_i T_1 = \sum_{j=1}^{n} a_{ij} \beta_j.
$$

Hence $\gamma_i R T_1 = \alpha_i T_1 = \sum_{j=1}^{n} a_{ij} \beta_j = \sum_{j=1}^{n} a_{ij} \delta_j S = \gamma_i T_2 S$,

so that $RT_1 = T_2 S$. $R$ and $S$ map a basis onto a basis, therefore they are non-singular. Hence $T_2 = RT_1 S^{-1}$.

Conversely, for the same choice of the pairs of bases and the matrices corresponding to the linear transformations relative to this choice of bases, we have $T_2 = RT_1 S^{-1}$ which implies $T_2 S = RT_1$. By previous calculations, $\gamma_i T_2 S = \gamma_i R T_1 = \sum_{j=1}^{n} a_{ij} \delta_j$. Hence $T_1$ and $T_2$ are represented by the same matrix $A$.

We now restrict our discussion to linear transformations of an $n$-dimensional space into itself. The matrices representing such transformations are square $n \times n$ matrices. If we set $V_m = W_n$ in Theorem 17, we have $\alpha = \beta$, $\gamma = \delta$, hence $R = S$ which implies that $P = Q$ or $Q^{-1} = P^{-1}$. Then the equivalence of matrices $A$ and $C$ takes the form $A = PCP^{-1}$. This special type of equivalence leads to still another definition.

**Definition.** Two $n \times n$ matrices $A$ and $B$ are similar if and only if $A = PBP^{-1}$ for some non-singular matrix $P$. 

Lemma 9. Similarity of matrices is an equivalence relation.

Proof. Let A, B, and C be $n \times n$ matrices. Let $A \sim B$ mean that $A$ is similar to $B$. Then $A = PBP^{-1}$ for a non-singular matrix $P$.

(i) For $P = I$, $A = IAI^{-1}$ so that $R$ is reflexive.

(ii) If $A \sim B$, then $A = PBF^{-1}$ where $P$ is non-singular. Multiplying the equation on the left by $P^{-1}$ and on the right by $P$, we obtain $B = P^{-1}AP$. There exists a non-singular matrix $Q$ such that $P^{-1} = Q$. Then $B = QAQ^{-1}$ so that $B \sim A$. Hence $R$ is symmetric.

(iii) If $A \sim B$ and $B \sim C$, then for suitable non-singular matrices $P$ and $Q$, we have $A = PBP^{-1}$ and $B = QCQ^{-1}$. Then $A = P(QCQ^{-1})P^{-1} = (PQ)C(PQ)^{-1}$ since multiplication of matrices is associative, and $Q^{-1}P^{-1} = (PQ)^{-1}$ by Theorem 4.9, (p. 78, Finkbeiner). Hence $A \sim C$. Therefore $R$ is transitive. Hence $R$ is an equivalence relation.

Theorem 19. Two square matrices are similar if and only if they represent the same linear transformation relative to suitably chosen bases.

Proof. Let $T$ be a linear transformation from $V_n$ into itself and let $\{\alpha_1, \ldots, \alpha_n\}$ and $\{\gamma_1, \ldots, \gamma_n\}$ be two
bases for $V_n$. With respect to the $\alpha$-basis, $T$ is represented uniquely by a matrix $A$. With respect to the $\gamma$-basis, $T$ is represented uniquely by a matrix $C$. Thus

$$\alpha_i T = \sum_{j=1}^{n} a_{ij} \alpha_j \text{ and } \gamma_j T = \sum_{k=1}^{n} c_{jk} \gamma_k, \quad i, j = 1, \ldots, n.$$ 

Let $R$ be a linear transformation which maps $\alpha_i$ onto $\gamma_i$. $R$ is non-singular since it maps a basis onto a basis, and relative to the $\gamma$-basis is represented by a non-singular matrix $P$ where $\alpha_i = \gamma_i R = \sum_{j=1}^{n} p_{ij} \gamma_j$. Calculating the image $\alpha T$ in two ways, we obtain

$$\alpha_i T = \sum_{j=1}^{n} a_{ij} \alpha_j = \sum_{j=1}^{n} a_{ij} (\gamma_j R) = \sum_{k=1}^{n} \left( \sum_{j=1}^{n} a_{ij} p_{jk} \right) \gamma_k,$$

$$\alpha_i T = (\gamma_i R) T = \left( \sum_{j=1}^{n} p_{ij} \gamma_j \right) T = \sum_{k=1}^{n} \left( \sum_{j=1}^{n} p_{ij} c_{jk} \right) \gamma_k,$$

for $i = 1, \ldots, n$.

Since $\alpha_i T$ is a unique linear combination of the $\gamma_k$,

$$\sum_{j=1}^{n} a_{ij} p_{jk} = \sum_{j=1}^{n} p_{ij} c_{jk}.$$

The equation represents the respective $(i,k)$ entries of $AP$ and $PC$, hence $AP = PC$. Since $P$ is non-singular ($P$ determines a non-singular transformation),

$$A = PCP^{-1}.$$ 

Hence, by definition, $A$ and $C$ are similar.

Conversely, if $A$ and $C$ are similar, then for a non-singular matrix $P$, $A = PCP^{-1}$ which implies $AP = PC$.

Choose any two bases as before, $\{\alpha\}$ and $\{\gamma\}$, and let the
matrices represent the linear transformations, as before, relative to these bases. Then we obtain the equation

\[ T^a_R = R T^y \]

corresponding to the above matrix equation. Hence

\[ Y_i^T R = (\sum_{j=1}^n c_{ij} y_j)^R = \sum_{j=1}^n c_{ij} (\sum_{k=1}^n p_{jk} y_k) \]

\[ = \sum_{k=1}^n (\sum_{j=1}^n c_{ij} p_{jk}) y_k = \alpha_i T \]

by previous calculation; \( Y_i^T R = (\gamma^T R) T = \alpha_i T \). Hence A and C determine the same \( T \) relative to the \( \alpha \)- and \( y \)-bases. This completes the proof.

Since the choice of a basis in a vector space determines a coordinate system for that space, the important consequence of Theorem 19 is that similar matrices represent the same linear transformation with respect to different coordinate systems. To put this in another way, similar matrices represent a linear transformation which is independent of a coordinate system of a given vector space.
V. INFINITE MATRICES

Preliminary remarks. In this section we discuss problems related to an extension of the results of the preceding sections to infinite matrices.

Consider a system of an infinity of homogeneous linear equations with real coefficients in an infinity of unknowns,

\[ \sum_{j=1}^{\infty} a_{ij} x_j = 0, \ i = 1, 2, \ldots \]  \(3\)

The coefficients determine an infinite matrix \( A = (a_{ij}) \), \( i, j = 1, 2, \ldots \). The sum of any two such matrices \( A + B = (a_{ij} + b_{ij}) \) always exists. In this section, let the scalars be real numbers. Multiplication by a scalar \( cA = (ca_{ij}) \) is always defined.

However, the matrix product \( AB = (c_{ij}) \) where

\[ c_{ij} = \sum_{k=1}^{\infty} a_{ik} b_{kj}, \ i, j = 1, 2, \ldots \], exists only if the sum converges. Consider an example: let \( A \) and \( B \) be infinite matrices

\[ A = \begin{pmatrix} 1 & 1 & 1 & \ldots \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \ldots \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \ldots \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{8} & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 1 & \ldots \\ 1 & 1 & 1 & \ldots \\ 1 & 1 & 1 & \ldots \\ 1 & 1 & 1 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \]

\( AB \) does not exist, since in each position of the \( i^{th} \) row, the element of \( AB \) is \( \sum_{k=1}^{\infty} \frac{1}{2^{i-1}} \), an infinite
sum of equal non-zero constants, hence diverges. However, BA does exist since in each position of the \( i \)th row of BA we have \( \sum_{i=1}^{\infty} \frac{1}{2^{i-1}} \), which is a convergent geometric series. This illustrates a problem of existence not encountered in our earlier work.

Cooke (1950, p. 8) shows that in general multiplication of infinite matrices is not associative. In the preceding sections, several results depended on the associative property of matrix multiplication.

In the finite case, the inverse of a matrix is unique; right and left inverses are the same. In general, this is not true for infinite matrices. In fact, Cooke shows (pp. 20-23) that numerous possibilities exist: an infinite matrix may have

(i) a unique two-sided inverse;
(ii) a unique two-sided inverse in a field \( F \) in which \( A^{-1}AA^{-1} \) is associative and infinitely many \( A^{-1}A \) or \( A^{-1} \) not belonging to \( F \) (where \( A^{-1} \) denotes a left inverse);
(iii) no inverse, either right or left;
(iv) no \( A^{-1} \) but infinitely many \( -1A \);
(v) no \( -1A \) but infinitely many \( A^{-1} \).

An example will illustrate case (v). Consider the matrices \( S \) and \( P \) where...
for arbitrary values in row 1 of \( P \). Hence \( P \) represents infinitely many right-inverses of \( S \). Now suppose that some \( Q \) is a left inverse of \( S \). Then the \((1,1)\) entry of \( QS \) is zero and the \((1,1)\) entry of \( I \) is 1. This contradiction proves that no left inverse of \( S \) exists.

In the preceding sections, in the results derived for a non-singular matrix \( A \), we assumed the existence and uniqueness of a two-sided inverse of \( A \).

The above remarks indicate that in considering an extension of the earlier results to infinite matrices, we must restrict our discussion to a subset with the following properties:

(i) matrix products always exist;

(ii) multiplication is associative;

(iii) if an inverse exists, it is unique and two-sided.

**Elementary matrices.** An elementary row operation can be performed also on an infinite matrix. The identity matrix \( I \) and the matrix \( E_{rs} \) we define as
before. Hence each type of elementary matrix, defined in terms of \( E_{rs} \), exists. We now relate the results in Section I to infinite matrices.

Consider a typical \((i,k)\) entry of \( IA, \sum_{j=1}^{\infty} \delta_{ij}a_{jk} \). \( \delta_{ij} = 0 \) if \( i \neq j \) and \( \delta_{ii} = 1 \), hence \( \sum_{j=1}^{\infty} \delta_{ij}a_{jk} = \delta_{ii}a_{ik} = a_{ik} \). This is the corresponding entry of \( A \). Hence \( IA = A \); in the same way, \( AI = A \). Thus \( IA = AI = A \) (which is Lemma 1) for any infinite matrix \( A \). Note that even though a sum such as \( \sum_{j=1}^{\infty} \delta_{ij}a_{jk} \) contains infinitely many terms, at most one term in the sum is non-zero, the rest are zeros.

Similar reasoning applies to the proofs of Theorems 1 and 3 and Lemma 3. Specifically, in the \((i,k)\) position, products of the type \( E_{ii}A, cE_{ii}A, \) and \( E_{ik}E_{hj} \) contain sums with infinitely many terms but at most one term in each sum is non-zero, owing to the special property of the infinite matrix \( E_{rs} \) which contains exactly one element 1 and the rest zeros. Thus Theorem 1, Lemma 3, and Theorem 3 extend to infinite matrices.

In Section I, we proved Lemma 2 and Theorem 2 by showing that the \( m \) rows of each \( m \times m \) matrix were linearly independent. Now linear independence is defined on a finite sum of terms. Hence infinitely many operations of taking a fixed number of terms each time may be required to prove that all the rows of an
infinite matrix are linearly independent. Therefore, to show that an infinite matrix is non-singular, a different method of proof is more convenient.

Lemma 2, of course, can be proved as follows: the $(i,k)$ entry of $II$ is $\sum_{j=1}^{\infty} \delta_{ij} \delta_{jk}$ which equals 1 if $i=j=k$ and zero otherwise. Hence $\sum_{j=1}^{\infty} \delta_{ij} \delta_{jk} = \delta_{ik}$, which is the corresponding entry of $I$. Hence $II = I$ so that $I$ is non-singular by definition. From the proof of Theorem 3 we have $P_{ij} P_{ij} = I$, $M_i(c) M_i(c^{-1}) = I$, and $A_{ij}(I - E_{jj}) = I$. It is easily verified that $M_i(c^{-1}) M_i(c) = I$ and $(I - E_{jj}) A_{ij} = I$. Thus every infinite elementary matrix has a two-sided inverse hence is non-singular by definition. Thus Theorem 2 is also extended. Therefore all the results of Section I extend to infinite matrices.

In the case of infinite elementary matrices multiplication is associative. For example,

\[
(P_{ij} M_i(c)) A_{ij} = I + E_{ij} - E_{jj} + c E_{ji} = P_{ij} (M_i(c) A_{ij}).
\]

Note. Hereafter, whenever the meaning is clear, by "matrix" we will mean "infinite matrix," and by "row operation" we will mean "elementary row operation."

Row equivalence. Recall that if $B$ is row equivalent to $A$, it can be obtained from $A$ through a finite number of row operations. We now consider the results of
Section II with respect to infinite matrices. Row equivalence is reflexive and transitive but it is symmetric only if in Lemma 5 the product $E_p^{-1} \ldots E_1^{-1} A$ exists.

Suppose that $B$ is row equivalent to $A$. If $B$ is non-singular, it contains infinitely many linearly independent row vectors since its rank is infinite. Hence we have no test for its relation to $A$ relative to rank. If, on the other hand, $B$ is singular, its rank is a finite number. In the latter case, the proof for Theorem 4 is similar.

An infinite matrix can be reduced to an echelon form by a finite number of row operations only if below a given number of rows, all the remaining rows have zero elements. For only in this case can induction be used. Under these conditions, Theorem 5 may be proved as before. If the matrix is not bounded below by zero rows, it would take infinitely many operations to reduce it to an echelon form.

Similarly, if $A$ is non-singular, we have no test for Theorem 6. But if beyond a given number of rows the remaining rows have all zeros, then $A$ is row equivalent to a reduced echelon matrix $E$ with $k$ non-zero rows, as before, hence $r(A) = k$. Then also $r(E') = k$, hence $r(A') = k$ if $A'E_1' \ldots E_k'$ exists.

Moreover, we cannot reduce an arbitrary non-singular
matrix to an identity matrix by a finite number of row operations; hence Theorems 7 and 8 cannot be extended by this method. Consequently, we cannot obtain an inverse of an arbitrary non-singular matrix by the methods of Section II.

If $B$ is non-singular, we can say nothing about Theorem 9. If, however, $r(B) = k$, say, then it is still true that $B = PA$ for a non-singular matrix $P$. But the converse does not extend, since in general $P$ is an infinite product of elementary matrices, which cannot be obtained by a finite number of row operations.

In general, Lemma 7 does not extend to the infinite case, since $A$ and $B$ both contain infinitely many rows, and a linear combination is defined as a sum of a finite number of terms. Also, row equivalence may not be symmetric. Hence Theorem 10 cannot be shown to hold.

Two matrices can be row equivalent and in reduced echelon form only if each has all zero rows beyond the first $k$ rows, say. Hence in general Theorem 11 cannot be shown to extend.

Thus in the infinite case row-equivalence is meaningful only for matrices with a finite number of non-zero rows.

**Equivalence.** Recall that if $B$ is equivalent to $A$,
it can be obtained from \( A \) by performing a finite number of elementary row and elementary column operations. We now consider the results of Section III relative to infinite matrices.

The converse of Theorem 12 does not extend since if \( P \) and \( Q \) are non-singular matrices, we cannot express them as finite products of elementary matrices.

Equivalence is reflexive but not necessarily symmetric since the product \( P^{-1}AQ^{-1} \) may not exist. Neither is it always transitive since \( P(\text{RCS})Q \) may not exist, and if it does exist, may not be associative.

Theorem 14 can be extended to infinite matrices only if the given matrix \( A \) has all zero rows beyond some finite number of rows, and all zero columns beyond a finite number of columns. For it is a matrix of this form only which can be reduced to the required form by a finite succession of row and column operations.

The first part of Theorem 15 can be established for matrices in the above form, since then the rank is a finite number. The converse of Theorem 15 holds only if the transitive property of Theorem 13 holds.

If an infinite matrix is non-singular, it may not be reducible to the identity matrix by a finite number of row and column operations. Hence it cannot, in general, be shown equivalent to the identity matrix by
by Theorem 16. In general, therefore, we cannot obtain an inverse of a non-singular matrix by the methods of Section III.

Hence in the infinite case equivalence is meaningful only for matrices with a finite number of non-zero rows and non-zero columns.

**Similarity.** Recall that if A is similar to B, then $A = PBP^{-1}$ for some non-singular matrix P. To consider similarity with respect to infinite matrices, $P^{-1}$ must be a unique inverse. But the uniqueness of $P^{-1}$ does not insure the existence of the product $P^{-1}AP$ in Lemma 9. Hence similarity is not necessarily symmetric. Clearly similarity is not always transitive as the product $PQCQ^{-1}P^{-1}$ may not exist. If it does exist and is associative, it may not be true that $Q^{-1}P^{-1} = (PQ)^{-1}$.

Theorem 19 (hence Theorem 17 of which this is a special case) in general may not extend to infinite matrices since $A = PBP^{-1}$ does not imply that AP exists.

To discuss this topic in greater depth would require an investigation of the properties of infinite-dimensional vector spaces and linear transformations on them.
BIBLIOGRAPHY
