# Equivalence of Matrices

Math 542

May 16, 2001

## 1 Introduction

The first thing taught in Math 340 is Gaussian Elimination, i.e. the process of transforming a matrix to reduced row echelon form by elementary row operations. Because this process has the effect of multiplying the matrix by an invertible matrix it has produces a new matrix for which the solution space of the corresponding linear system is unchanged. This is made precise by Theorem 2.4 below.

The theory of Gaussian elimination has the following features:

1. There is an equivalence relation which respects the essential properties of some class of problems. Here the equivalence relation is called *row equivalence* by most authors; we call it *left equivalence*.

2. The equivalence classes of this relation are the orbits of a group action. In the case of left equivalence the group is the general linear group acting by left multiplication.

3. There is a characterization of the equivalence relation in terms of some invariant (or invariants) associated to a matrix. In the case of left equivalence the characterization is provided by Theorem 2.4 which says that two matrices of the same size are left equivalent if and only if they have the same null space.

4. There is a normal form and a theorem which says that each matrix is equivalent to a unique matrix in normal form. In the case of left equivalence the normal form is *reduced row echelon form* (not explained in this paper).

Our aim in this paper is to give other examples of equivalence relations which fit this pattern. Two examples (*right [column] equivalence* and *left right equivalence*) are (or should be) standard parts of the undergraduate curriculum; two others (*lower equivalence* and *lower upper equivalence*) are not as well known but not appreciably more difficult.

This paper is the result of a term paper I assigned in my Math 542 class in the spring semester of 2001 at the University of Wisconsin. I provided the students with an outline, criticized their first draft, and wrote this paper to fulfill a promise that they would have my version of the term paper when they handed in their final draft. You can generate the outline by modifying the LaTeX file that produced this document: replace the command `\tellfalse` by `\telltrue`.

# 2 Statement of the Theorems

**2.1.** Throughout $\mathbb{F}$ is a field and $\mathbb{F}^{m\times n}$ is the vector space (over $\mathbb{F}$) of all $m\times n$ matrices with entries from $\mathbb{F}$. We write $\mathbb{F}^n$ instead of $\mathbb{F}^{n\times 1}$. A square matrix is **diagonal** iff all the entries off the diagonal vanish, **upper triangular** iff all the entries below the diagonal vanish, **lower triangular** iff all the entries above the diagonal vanish, and **unitriangular** iff it is upper triangular and its diagonal entries are one. In other words, a matrix

- $P$ is diagonal iff $\text{entry}_{ij}(P) = 0$ for $i \neq j$,
- $P$ is upper triangular iff $\text{entry}_{ij}(P) = 0$ for $i > j$,
- $Q$ is lower triangular iff $\text{entry}_{ij}(Q) = 0$ for $i < j$, and
- $P$ is unitriangular iff $\text{entry}_{ij}(P) = 0$ for $i > j$ and $\text{entry}_{ii}(P) = 1$.

**2.2.** The **general linear group** in dimension $n$ is the set $\mathbb{GL}_n(\mathbb{F})$ of all invertible $n \times n$ matrices with entries from $F$. Define

$$\begin{aligned}
\mathbb{D}_n(\mathbb{F}) &= \{P \in \mathbb{GL}_n(\mathbb{F}) : P \text{ is diagonal}\},\\
\mathbb{B}_n(\mathbb{F}) &= \{P \in \mathbb{GL}_n(\mathbb{F}) : P \text{ is upper triangular}\},\\
\mathbb{B}'_n(\mathbb{F}) &= \{P \in \mathbb{GL}_n(\mathbb{F}) : P \text{ is lower triangular}\},\\
\mathbb{U}_n(\mathbb{F}) &= \{P \in \mathbb{GL}_n(\mathbb{F}) : P \text{ is unitriangular}\}.
\end{aligned}$$

**2.3. Definition.** Let $A, B \in \mathbb{F}^{m\times n}$ be two matrices of the same size. We say that

- $A$ is **left equivalent** to $B$ iff there exists $Q \in \mathbb{GL}_m(\mathbb{F})$ such that
$$A = QB.$$
- $A$ is **right equivalent** to $B$ iff there exists $P \in \mathbb{GL}_n(\mathbb{F})$ such that
$$A = BP^{-1}.$$
- $A$ is **right left equivalent** to $B$ iff there exist $Q \in \mathbb{GL}_m(\mathbb{F})$ and $P \in \mathbb{GL}_n(\mathbb{F})$ such that
$$A = QBP^{-1}.$$
- $A$ is **lower equivalent** to $B$ iff there exists $Q \in \mathbb{B}'_m(\mathbb{F})$ such that
$$A = QB.$$
- $A$ is **lower upper equivalent** to $B$ iff there exist $Q \in \mathbb{B}'_m(\mathbb{F})$ and $P \in \mathbb{U}_n(\mathbb{F})$ such that
$$A = QBP^{-1}.$$

These are all equivalence relations, i.e. where $A \equiv B$ denotes any of these four relations we have

$$A \equiv A; \qquad \text{(Reflexive Law)}$$
$$A \equiv B \implies B \equiv A; \qquad \text{(Symmetric Law)}$$
$$A \equiv B,\ B \equiv C \implies A \equiv C. \qquad \text{(Transitive Law)}$$

This follows immediately from the fact that $\mathbb{GL}_m(\mathbb{F})$ and $\mathbb{GL}_n(\mathbb{F})$ are groups the sets $\mathbb{B}'_m(\mathbb{F}) \subset \mathbb{GL}_m(\mathbb{F})$ and $\mathbb{U}_n(\mathbb{F}) \subset \mathbb{GL}_n(\mathbb{F})$ are subgroups. (See Corollary 3.5 below.)

**Theorem 2.4.** *Suppose that $A, B \in \mathbb{F}^{m\times n}$. Then $A$ and $B$ are left equivalent if and only if they have the same null space.*

**Theorem 2.5.** *Suppose that $A, B \in \mathbb{F}^{m\times n}$. Then $A$ and $B$ are right equivalent if and only if they have the same range.*

**Theorem 2.6.** *Suppose that $A, B \in \mathbb{F}^{m\times n}$. Then $A$ and $B$ are right left equivalent if and only if they have the same rank.*

**2.7. Definition.** For $A \in \mathbb{F}^{m\times n}$, $p = 1, \ldots, m$, $q = 1, \ldots, n$, let $A_{pq} \in \mathbb{F}^{p\times q}$ denote the $p \times q$ matrix which forms upper left hand corner of $A$ so that

$$\mathrm{entry}_{ij}(A_{pq}) = \mathrm{entry}_{ij}(A)$$

for $i = 1, \ldots, p$, $j = 1, \ldots, q$. The $(p, q)$ **corner rank** of $A$ is defined by

$$\delta_{pq}(A) = \mathrm{rank}(A_{pq}).$$

**Theorem 2.8.** *Suppose that $A, B \in \mathbb{F}^{m\times n}$. Then $A$ and $B$ are lower upper equivalent if and only if they have the same corner ranks.*

**2.9. Definition.** Let $W$ be a vector space of dimension $m$ over the field $\mathbb{F}$. A **flag** in $W$ is a sequence of subspaces

$$\{0\} = W_0 \subset W_1 \subset W_2 \subset \cdots \subset W_m = W$$

such that $\dim(W_k) = k$ for $k = 0, 1, 2, \ldots, m$. The **standard flag** in $\mathbb{F}^m$ is defined by

$$W_k = \mathrm{Span}(e_1, e_2, \ldots, e_k)$$

where $(e_1, e_2, \ldots, e_m)$ is the **standard basis**, i.e. $e_i$ is the $i$th column of the identity matrix $I_m$. The **reverse standard flag** in $\mathbb{F}^m$ is defined by

$$W'_k = \mathrm{Span}(e_{m-k+1}, e_{m-k+2}, \ldots, e_m).$$

**Theorem 2.10.** *Suppose that $A, B \in \mathbb{F}^{m\times n}$. Then $A$ and $B$ are lower equivalent if and only if*

$$A^{-1}(W'_k) = B^{-1}(W'_k)$$

*for $k = 0, 1, 2, \ldots, m$ where $W'_0, W'_1, \ldots, W'_m$ is the reverse standard flag in $\mathbb{F}^m$.*

Theorems 2.4, 2.5, and 2.6 are proved in Section 4; Theorem 2.8 is proved in Section 5; and Theorem 2.10 is proved in Section 6.

**2.11. Remark.** Theorem 2.10 is Exercise 345E in [2]. The hint given there is only appropriate when the matrices $A$ and $B$ have rank $m$.

## 3 Elementary Matrices

**3.1.** Recall from Math 340 that there are three kinds of **elementary matrices** as follows.

**Scale** The elementary matrix $\mathrm{Scale}(I_m, p, c)$ results from the $m \times m$ identity matrix $I_m$ by multiplying the $p$th row by $c$.

**Swap** The elementary matrix $\mathrm{Swap}(I_m, p, q)$ results from the $m \times m$ identity matrix $I_m$ by interchanging rows $p$ and $q$.

**Shear** The elementary matrix $\mathrm{Shear}(I_m, p, q, c)$ results from the $m \times m$ identity matrix $I_m$ by adding $c$ times the $q$th row to the $p$th row.

Shears and swaps are defined only if $p \neq q$. The *Fundamental Theorem on Row Operations* (see [1] page 54) says that the matrix which results by multiplying a matrix $A \in \mathbb{F}^{m \times n}$ on the left by an elementary matrix is the same as the matrix which results by applying the corresponding elementary row operation to $A$, i.e.

- $E = \mathrm{Scale}(I_m, p, c) \implies EA$ results from $A$ by multiplying the $p$th row by $c$;
- $E = \mathrm{Swap}(I_m, p, q) \implies EA$ results from $A$ by interchanging rows $p$ and $q$;
- $E = \mathrm{Shear}(I_m, p, q, c) \implies EA$ results from $A$ by adding $c$ times the $q$th row to the $p$th row.

The *Fundamental Theorem on Column Operations* says that multiplying on the right performs the corresponding column operation, i.e.

- $E = \mathrm{Scale}(I_n, p, c) \implies AE$ results from $A$ by multiplying the $p$th column by $c$;
- $E = \mathrm{Swap}(I_n, p, q) \implies AE$ results from $A$ by interchanging columns $p$ and $q$;
- $E = \mathrm{Shear}(I_n, p, q, c) \implies AE$ results from $A$ by adding $c$ times the $p$th column to the $q$th column.

An **upper shear** is an upper triangular shear matrix. A **lower shear** is a lower triangular shear matrix.

**3.2.** Elementary matrices are invertible; in fact, the inverse of an elementary matrix is an elementary matrix of the same type:

- $E = \operatorname{Scale}(I_n, p, c) \implies E^{-1} = \operatorname{Scale}(I_n, p, 1/c)$;
- $E = \operatorname{Swap}(I_n, p, q) \implies E^{-1} = E$;
- $E = \operatorname{Shear}(I_n, p, q, c) \implies E^{-1} = \operatorname{Shear}(I_n, p, q, -c)$

**3.3.** A set $S$ of invertible matrices is said to **generate** a group $G$ of invertible matrices iff (1) $S \subseteq G$, and (2) every element of $G$ is the product of a finite number of elements of $S$. It is an easy consequence of the Fundamental Theorem that

**Theorem.** *The elementary matrices generate* $\mathbb{GL}_n(\mathbb{F})$.

For the proof see [1] Page 59 for example, or modify the arguments described below. The following theorem is a refinement.

**3.4. Factorization Theorem.**

**(i)** *The invertible scales generate* $\mathbb{D}_n(\mathbb{F})$.

**(ii)** *The upper shears generate* $\mathbb{U}_n(\mathbb{F})$.

**(iii)** *The invertible scales and upper shears generate* $\mathbb{B}_n(\mathbb{F})$.

**(iv)** *The invertible scales and lower shears generate* $\mathbb{B}'_n(\mathbb{F})$.

*Proof.* An element $D \in \mathbb{D}_n(F)$ is a product $D = E_1 E_2 \cdots E_n$ of elementary scales where $E_i = \operatorname{Scale}(I_n, i, d_i)$ and $d_i$ is the $i$th diagonal entry of $D$. This proves (i).

To prove (ii) note first that if $U \in U_n(\mathbb{F})$ and $E$ is an upper shear then $EU \in U_n(\mathbb{F})$. This follows from the Fundamental Theorem; subtracting a row of $U$ from a row above leaves entries to the left of the diagonal unchanged. Now let $E_{p,q} = \operatorname{Shear}(I_n, p, q, -u_{p,q})$. Then, by the Fundamental Theorem, the matrix $E_{1,n} E_{2,n} \cdots E_{n-1,n} U$ agrees with $U$ except in the $n$th column which is replaced by the $n$th column of the identity matrix $I_n$. Repeat this process for columns $n-1, n-2, \ldots, 3, 2$ (in that order) and we obtain

$$E_{1,2}(E_{1,3}E_{2,3})(E_{1,4}E_{2.4}E_{3,4}) \cdots (E_{1,n}E_{2,n} \cdots E_{n-1,n})U = I_n$$

which (as the inverse of an elementary matrix is an elementary matrix of the same type) proves (ii). In case $n = 3$ the above factorization takes the form

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Item (iii) follows from (i) and (ii) as every element of $\mathbb{B}_n(\mathbb{F})$ is the product of an element of $\mathbb{D}_n(\mathbb{F})$ and an element of $\mathbb{U}_n(\mathbb{F})$. Item (iv) follows from item (iii) as the transpose of an element of $\mathbb{B}_n(\mathbb{F})$ is an element of $\mathbb{B}'_n(\mathbb{F})$ (and vice versa) and the transpose of an upper shear is a lower shear. □

**Corollary 3.5.** *The sets $\mathbb{D}_n(\mathbb{F})$, $\mathbb{B}_n(\mathbb{F})$, $\mathbb{B}'_n(\mathbb{F})$, and $\mathbb{U}_n(\mathbb{F})$ are all subgroups of the group $\mathbb{GL}_n(\mathbb{F})$.*

*Proof.* It is easy to see that each of these sets contains the identity matrix and is closed under multiplication. That these sets are closed under the inverse operation follows from the Factorization Theorem and the fact that the inverse of an elementary matrix is an elementary matrix of the same type. □

# 4 Null Space, Range, and Rank

In this section we prove Theorems 2.4, 2.5, and 2.6. Recall that the **null space** of a matrix $A \in \mathbb{F}^{m\times n}$ is the subspace

$$\mathcal{N}(A) = \{v \in \mathbb{F}^n : Av = 0\},$$

the **range** of $A$ is the subspace

$$\mathcal{R}(A) = \{Av : v \in \mathbb{F}^n\},$$

the **nullity** of $A$ is the dimension of $\mathcal{N}(A)$, and the **rank** of $A$ is the dimension of $\mathcal{R}(A)$.

**Lemma 4.1.** *Suppose that $\gamma_1, \ldots \gamma_{r+k} \in \mathbb{F}^n$ satisfy*

**(i)** *$A\gamma_1, \ldots, A\gamma_r$ form a basis for $\mathcal{R}(A)$, and*

**(ii)** *$\gamma_{r+1}, \ldots \gamma_{r+k}$ form a basis for $\mathcal{N}(A)$.*

*Then $\gamma_1, \dots \gamma_{r+k}$ forms a basis for $\mathbb{F}^n$.*

*Proof. We show $\gamma_1, \dots \gamma_{r+k}$ are linearly independent.* Assume $c_1, \dots, c_{r+k} \in \mathbb{F}$ satisfy

$$c_1\gamma_1 + \cdots + c_r\gamma_r + c_{r+1}\gamma_{r+1} + \cdots + c_{r+k}\gamma_{r+k} = 0. \tag{1}$$

Multiply by $A$: as $\gamma_{r+1}, \dots \gamma_{r+k} \in \mathcal{N}(A)$ we obtain

$$c_1A\gamma_1 + \cdots + c_rA\gamma_r = 0.$$

By (i) $c_1 = \cdots = c_r = 0$. Hence (1) becomes $c_{r+1}\gamma_{r+1} + \cdots + c_{r+k}\gamma_{r+k} = 0$ so by (ii) we get $c_{r+1} = \cdots = c_{r+k} = 0$.

*We show $\gamma_1, \dots \gamma_{r+k}$ span $\mathbb{F}^n$.* Choose $v \in \mathbb{F}^n$. Then $Av \in \mathcal{R}(A)$ so by (i) there exist $c_1, \dots, c_r \in \mathbb{F}$ such that

$$Av = c_1A\gamma_1 + \cdots c_rA\gamma_r.$$

Hence $v - (c_1\gamma_1 + \cdots + c_r\gamma_r) \in \mathcal{N}(A)$ so by (ii) there exist $c_{r+1}, \dots, c_{r+k} \in \mathbb{F}$ with

$$v - (c_1\gamma_1 + \cdots + c_r\gamma_r) = c_{r+1}\gamma_{r+1} + \cdots + c_{r+k}\gamma_{r+k}$$

so $v$ is a linear combination of $\gamma_1, \dots, \gamma_{r+k}$ as required. □

**Corollary 4.2 (Rank Nullity Relation).** *For $A \in \mathbb{F}^{m\times n}$ the rank of $A$ plus the nullity of $A$ is $n$.*

*Proof of 2.4.* Assume that $A$ and $B$ are left equivalent, i.e. that $A = QB$ where $Q \in \mathbb{GL}_m(\mathbb{F})$. Then as $Q$ is invertible we have

$$Av = 0 \iff QBv = 0 \iff Bv = 0$$

Therefore $\mathcal{N}(A) = \mathcal{N}(B)$ as required.

Conversely, assume that $\mathcal{N}(A) = \mathcal{N}(B)$. Let $\gamma_{r+1}, \dots, \gamma_n$ be a basis for $\mathcal{N}(A) = \mathcal{N}(B)$ and extend to a basis $\gamma_1, \dots, \gamma_n$. Then $A\gamma_1, \dots, A\gamma_r$ are linearly independent for otherwise some linear combination of $\gamma_1, \dots, \gamma_r$ would lie in the null space of $A$ contradicting the independence of $\gamma_1, \dots, \gamma_n$. Let $\alpha_i = A\gamma_i$ for $i = 1, \dots, r$ and extend to a basis $\alpha_1, \dots \alpha_m$ of $\mathbb{F}^m$. Similarly there is a basis $\beta_1, \dots \beta_m$ of $\mathbb{F}^m$ such that $\beta_i = B\gamma_i$ for $i = 1, \dots, r$. Define $Q$ by $Q\beta_i = \alpha_i$ for $i = 1, \dots, m$. Then for $i = 1, \dots, r$ we have

$$A\gamma_i = \alpha_i = Q\beta_i = QB\gamma_i.$$

This also holds for $i = r+1, \dots, n$ as both sides are zero. Hence $A = QB$ so $A$ and $B$ are left equivalent as required. □

*Proof of 2.5.* Assume that $A$ and $B$ are right equivalent, i.e. that $A = BP^{-1}$ where $P \in \mathbb{GL}_n(\mathbb{F})$. Then $P(F^n) = \mathbb{F}^n$ as $P$ is invertible, so

$$\mathcal{R}(A) = A(\mathbb{F}^n) = A(P(\mathbb{F}^n)) = AP(F^n) = B(F^n) = \mathcal{R}(B)$$

as required.

Conversely assume $\mathcal{R}(A) = \mathcal{R}(B)$. Let $\phi_1, \ldots, \phi_r$ be a basis for this space and choose $\gamma_1, \ldots, \gamma_r$ and $\gamma'_1, \ldots, \gamma'_r$ so that

$$A\gamma_i = \phi_i = B\gamma'_i \tag{2}$$

for $i = 1, \ldots, r$. Let $\gamma_{r+1}, \ldots \gamma_n$ be a basis for $\mathcal{N}(A)$ and $\gamma'_{r+1}, \ldots \gamma'_n$ be a basis for $\mathcal{N}(B)$. By 4.1 $\gamma_1, \ldots, \gamma_n$ is a basis for $\mathbb{F}^n$ as is $\gamma'_1, \ldots, \gamma'_n$. Then (2) holds for $i = r+1, \ldots, n$ since both sides are zero. Define $P$ by $\gamma_i = P\gamma'_i$ for $i = 1, 2, \ldots, n$. Then $AP\gamma'_i = A\gamma_i = B\gamma'_i$ for $i = 1, 2, \ldots, n$ so $AP = B$ as required. □

*Proof of 2.6.* Assume that $A$ and $B$ are left right equivalent, i.e. that $A = QBP^{-1}$ where $Q \in \mathbb{GL}_m(\mathbb{F})$ and $P \in \mathbb{GL}_n(\mathbb{F})$. Then

$$\mathcal{R}(A) = A(\mathbb{F}^n) = QBP^{-1}(\mathbb{F}^n) = QB(\mathbb{F}^n) = Q\mathcal{R}(B),$$

i.e. the isomorphism $Q : \mathbb{F}^m \to \mathbb{F}^m$ restricts to an isomorphism from $\mathcal{R}(A)$ to $\mathcal{R}(B)$. Hence

$$\operatorname{rank}(A) = \dim \mathcal{R}(A) = \dim \mathcal{R}(B) = \operatorname{rank}(B)$$

as required.

For the converse we introduce the matrix

$$D_{m,n,r} = \begin{bmatrix} I_r & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix}. \tag{3}$$

It is easy to see that for any matrix $A \in \mathbb{F}^{m \times n}$ there is a basis $\gamma_1, \ldots, \gamma_n$ as in Lemma 4.1. Moreover if $\gamma_1, \ldots, \gamma_n$ are the columns of the identity matrix $I_n$ and $A\gamma_1, \ldots, A\gamma_r$ are the first $r$ columns of the the identity matrix $I_m$ then $A = D_{m,n,r}$. It follows as in the proofs of 2.4 and 2.5 that that any matrix is right left equivalent to some $D_{m,n,r}$. But the rank of $D_{m,n,r}$ is $r$. Hence if $A$ and $B$ have the same rank $r$ they are each right left equivalent to $D_{m,n,r}$ and hence to each other. □

# 5 Rook Matrices

In this section we prove Theorem 2.8.

**5.1.** A matrix is called a **rook matrix**, iff all its entries are either 0 or 1 and it has at most one nonzero entry in every row and at most one nonzero entry in every column. An invertible rook matrix is also called a **permutation matrix**. The matrix $D_{m,n,r} \in \mathbb{F}^{m\times n}$ defined by (3) above is an example of a rook matrix.

**5.2. Remark.** The $n \times n$ permutation matrices form a finite group isomorphic to the group $S_n$ of permutations of the finite set $\{1, 2, \ldots, n\}$. Just as the transpositions generate the latter, the swap matrices generate the former.

In Math 340 it is proved that every matrix is left equivalent to a matrix $R$ in reduced row echelon form; one can show that this matrix $R$ is unique. The following theorem is analogous. (Another analog is Corollary 5.9 below.)

**Theorem 5.3 (Rook Decomposition).** *Every matrix is lower upper equivalent to a unique rook matrix.*

**Lemma 5.4.** *If $Q \in \mathbb{B}'_m(F)$ then $Q_{pp} \in \mathbb{B}'_p(\mathbb{F})$ for $p = 1, 2, \ldots, m$. Similarly, If $P \in \mathbb{B}_n(F)$ then $P_{qq} \in \mathbb{B}_q(\mathbb{F})$ for $q = 1, 2, \ldots, n$.*

*Proof.* It is clear that $Q_{pp}$ is triangular as it is the upper left hand corner of a triangular matrix. A triangular matrix is invertible if and only if its diagonal entries are nonzero; hence the fact that $Q$ is invertible implies that $Q_{pp}$ is. The proof for $P$ is the same. □

**Lemma 5.5.** *Lower upper equivalent matrices have the same corner ranks.*

*Proof.* Assume that $AP = QB$ where $Q \in \mathbb{B}'_m(\mathbb{F})$ and $P \in \mathbb{B}_n(\mathbb{F})$. Then

$$AP = \begin{bmatrix} A_{pq} & * \\ * & * \end{bmatrix} \begin{bmatrix} P_{qq} & * \\ 0 & * \end{bmatrix} = \begin{bmatrix} A_{pq}P_{qq} & * \\ * & * \end{bmatrix}$$

and

$$QB = \begin{bmatrix} Q_{pp} & 0 \\ * & * \end{bmatrix} \begin{bmatrix} B_{pq} & * \\ * & * \end{bmatrix} = \begin{bmatrix} Q_{pp}B_{pq} & * \\ * & * \end{bmatrix}$$

so $A_{pq}P_{qq} = Q_{pp}B_{pq}$ so $A_{pq} = Q_{pp}B_{pq}P_{qq}^{-1}$ so

$$\delta_{pq}(A) = \text{rank}(A_{pq}) = \text{rank}(B_{pq}) = \delta_{pq}(B)$$

as required. □

**Lemma 5.6.** *For a matrix $D$ in rook normal form the corner ranks are given by*

$$\delta_{pq}(D) = \sum_{i=1}^{p} \sum_{j=1}^{q} \mathrm{entry}_{ij}(D).$$

*Proof.* The sum on the right is the number of nonzero columns of $D_{pq}$. These columns are independent because they are distinct columns of the identity matrix $I_p$. □

**Corollary 5.7.** *Two matrices in rook normal form are equal if and only if they have the same corner ranks.*

*Proof.* By Lemma 5.6 a rook matrix $D$ satisfies

$$\delta_{pq}(D) = \delta_{p-1,q}(D) + \delta_{p,q-1}(D) - \delta_{p-1,q-1}(D) + \mathrm{entry}_{pq}(D).$$

□

*Proof of 5.3.* The theorem says that every matrix $A$ can be transformed to a rook matrix by elementary row and column operations where the only row operations allowed are scales and lower shears and the only column operations allowed are upper shears and that the resulting rook matrix is independent of the order of the operations. Figure 1 gives an algorithm for doing this. Uniqueness is proved as follows. If $A$ is lower upper equivalent to rook matrices $D$ and $D'$ then, by Lemma 5.5, $A$ and $D$ have the same corner ranks as do $A$ and $D'$. Hence $D = D'$ by Lemma 5.7.

□

**Corollary 5.8 (Bruhat Decomposition).** *Every invertible matrix is lower upper equivalent to a unique permutation matrix.*

*Proof.* This is a special case of: see 5.1. □

*Proof of 2.8.* 'Only if' is Lemma 5.5. For the converse assume that $A, B \in \mathbb{F}^{m\times n}$ have the same corner ranks. By Theorem 5.3 there are rook matrices $D$ and $D'$ with $A$ right left equivalent to $D$ and $B$ right left equivalent to $D'$. By Lemma 5.5 $D$ and $D'$ have the same corner ranks, so by Lemma 5.7 $D = D'$. Hence $A$ is right left equivalent to $B$. □

**Corollary 5.9.** *Every $A \in \mathbb{F}^{m\times n}$ is left right equivalent to a unique matrix $D_{m,n,r}$.*

```
for q=1:n %  loop on columns
  for p=1:m
     if (A(p,q)!=0)
         A(p,:) = A(p,:)/A(p,q)      % scale
         for i=p+1:m                 % shear
             A(i,:) = A(i,:) - A(i,q)*A(p,:)
         end
         go to next_Col
     end
  end next_Col
end
for p=1:m  %  loop on rows
    q=1;
    while (A(p,q)==0) q=q+1  end
    for j=q+1:n                      % shear
       A(:,j) = A(:,j) - A(p,j)*A(:,q)
    end
end
```

Figure 1: Computing the rook matrix

*Proof.* This was proved in the proof of 2.6. We can also deduce it from 5.3 as follows. According to the Fundamental Theorem left (resp. right) multiplication by a permutation matrix permutes the rows (resp. columns) accordingly. More precisely, for each permutation $\sigma \in S_n$ in the symmetric group $S_n$ on $n$ symbols, the permutation matrix $Q$ determined by

$$\mathrm{row}_i(Q) = \mathrm{row}_{\sigma(i)}(I_m)$$

for $i = 1, 2, \ldots, m$ satisfies

$$\mathrm{row}_i(QA) = \mathrm{row}_{\sigma(i)}(A)$$

for $A \in \mathbb{F}^{m \times n}$. A similar statement holds for columns; namely if

$$\mathrm{col}_j(P) = \mathrm{col}_{\tau(j)}(I_m)$$

for $i = 1, 2, \ldots, m$ and $\tau \in S_m$, then

$$\mathrm{col}_j(AP) = \mathrm{col}_{\tau(j)}(A).$$

Now it is clear that for any rook matrix $D$ there are permutation matrices $Q$ and $P$ such that $QDP^{-1} = D_{m,n,r}$ which proves existence in 5.9. Uniqueness follows from 2.6 and the fact that the rank of $D_{m,n,r}$ is $r$. $\square$

# 6 Flags

In this section we prove Theorem 2.10. As a warmup, consider the case $n = 1$ and $m = 2$. Then

$$A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \qquad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

and a typical element $Q \in \mathbb{B}_2'(\mathbb{F})$ has form

$$Q = \begin{bmatrix} q_1 & 0 \\ q_2 & q_3 \end{bmatrix}$$

where $q_1 q_3 \neq 0$. Then $A^{-1}(W_2') = \{0\}$ if and only if $a_1 \neq 0$ and the equation $A = QB$ can be solved for $Q \in \mathbb{B}_2'(\mathbb{F})$ if and only if either $a_1 b_1 \neq 0$ or $a_1 = b_1 = 0$ and $a_2 b_2 \neq 0$ or $A = B = 0$.

**Lemma 6.1.** *Suppose that $Q \in \mathbb{GL}_m(\mathbb{F})$ is invertible. Then*

$$Q \in \mathbb{B}_m(\mathbb{F}) \iff QW_k = W_k \text{ for } k = 1, 2, \ldots, m$$

*where $W_0, \ldots, W_m$ is the standard flag. Similarly*

$$Q \in \mathbb{B}'_m(\mathbb{F}) \iff QW'_k = W'_k \text{ for } k = 1, 2, \ldots, m$$

*for $k = 1, 2, \ldots, m$ where $W'_0, \ldots, W'_m$ is the reverse standard flag.*

*Proof.* Recall that $W_k = \mathrm{Span}(e_1, \ldots, e_k)$ where $e_i$ is the $i$th column of the identity matrix. For any matrix $Q$ we have

$$Qe_k = \sum_{i=1}^{m} \mathrm{entry}_{ik}(Q) e_i$$

and $Q$ is upper triangular if and only if

$$Qe_k = \sum_{i=1}^{k} \mathrm{entry}_{ik}(Q) e_i,$$

for all $k$ i.e. if and only if $Qe_k \in W_k$ for all $k$. The proof in the lower triangular case is essentially the same. □

*Proof of 2.8.* Assume that $A = QB$ where $Q \in \mathbb{B}'_m(F)$. Then by Lemma 6.1 we have

$$A^{-1}(W'_k) = B^{-1}(Q^{-1}(W'_k)) = B^{-1}(W'_k)$$

as required.

Conversely assume the preimages by $A$ and $B$ of the subspace $W_k$ are equal and denote this preimage by $V_k$. Thus

$$V_k = A^{-1}(W'_k) = B^{-1}(W'_k)$$

and

$$\mathcal{N}(A) = \mathcal{N}(B) = V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots \subseteq V_m = \mathbb{F}^{n \times 1}.$$

Let

$$n_k = \dim(V_k).$$

**Claim.** There are bases $(\alpha_1 \ldots, \alpha_m)$ and $(\beta_1, \ldots, \beta_m)$ of $\mathbb{F}^{m \times 1}$ and a basis $(\gamma_1, \ldots, \gamma_n)$ of $\mathbb{F}^{n \times 1}$ such that for $k = 0, 1, \ldots, m$

**(i)** $W'_k = \text{Span}(\alpha_1, \ldots, \alpha_k) = \text{Span}(\beta_1, \ldots, \beta_k)$;

**(ii)** $(\gamma_1, \ldots, \gamma_{n_k})$ is a basis for $V_k$; and

**(iii)** $A\gamma_{n_k} = \alpha_k$ and $B\gamma_{n_k} = \beta_k$ if $n_{k-1} < n_k$.

To prove the claim take $(\gamma_1, \ldots, \gamma_{n_0})$ to be any basis for the common null space $V_0$ of $A$ and $B$. Now proceed inductively: assume that $\alpha_i$ and $\beta_i$ have been defined for $i \leq k-1$ and that $\gamma_j$ has been defined for $j \leq n_{k-1}$.

**Subclaim.** Either $n_k = n_{k-1}$ or $n_k = n_{k-1} + 1$. Indeed, otherwise there are vectors $v_1$ and $v_2$ in $V_k$ such that $\gamma_1, \ldots, \gamma_{n_{k-1}}, v_1, v_2$ are independent. Now $Av_1$ and $Av_2$ lie in $W'_k$ so some linear combination $c_1Av_1 + c_2Av_2$ lies in $W'_{k-1}$. But then $c_1v_1 + c_2v_2$ lies in $W'_{k-1}$ and is thus a linear combination of $\gamma_{n_1}, \ldots, \gamma_{n_{k-1}}$ contradicting the assumption that $\gamma_1, \ldots, \gamma_{n_{k-1}}, v_1, v_2$ are independent. This proves the subclaim.

Now if $n_{k-1} = n_k$ let $\alpha_k = \beta_k = e_k$ the $(n-k+1)$st column of the $m \times m$ identity matrix. Otherwise extend the basis $\gamma_1, \ldots, \gamma_{n_{k-1}}$ of $V_{k-1}$ to a basis $(\gamma_1, \ldots, \gamma_{n_k})$ of $V_k$ and define $\alpha_k = A\gamma_{n_k}$ and $\beta_k = B\gamma_{n_k}$. In either case $\alpha_k$ and $\beta_k$ are in $W'_k$ but not in $W'_{k-1}$. Since $(\alpha_1, \ldots, \alpha_{k-1})$ and $(\beta_1, \ldots, \beta_{k-1})$ are bases of $W'_{k-1}$ it follows that $(\alpha_1, \ldots, \alpha_k)$ and $(\beta_1, \ldots, \beta_k)$ are bases of $W'_k$. This proves the claim.

Since $(\beta_1, \ldots, \beta_m)$ is a basis for $\mathbb{F}^{m\times 1}$ there is a unique matrix $Q$ such that $Q\beta_k = \alpha_k$ for $k = 1, 2, \ldots, m$, and since $(\alpha_1 \ldots, \alpha_m)$ is also a basis, the matrix $Q$ is invertible. By (i) and Lemma 6.1 $Q \in \mathbb{B}'_m(\mathbb{F})$. By (iii) we have $QB\gamma_j = A\gamma_j$ for $j = n_1, \ldots, n_m = n$ and by (ii) (with $k = 0$) we have $QB\gamma_j = A\gamma_j = 0$ for $j = 1, \ldots, n_0$. Hence $QB\gamma_j = A\gamma_j$ for all $j$. By (ii) (with $k = m$) the sequence $(\gamma_1, \ldots, \gamma_n)$ is a basis for $\mathbb{F}^{n\times 1}$ so $QB = A$. □

# References


[1] Steven J. Leon: *Linear Algebra with Applications (4th Edition)*, Prentice Hall, 1994.

[2] J. W. Robbin: *Matrix Algebra Using MINImal MATlab*, A.K. Peters, 1995.