3 Elementary Matrix Operations and Systems of Linear Equations

3.1 Elementary Matrix Operations and Elementary Matrices

In this chapter we develop a (hopefully!) familiar method for comparing matrices.

Definition 3.1. An *elementary row operation* is one of three transformations of the rows of a matrix:

Type I: Swap two rows;

Type II: Multiply a row by a non-zero constant;

Type III: Add to one row a scalar multiple of another.

Matrices are *row equivalent* if there exists a finite sequence of elementary row operations transforming one to the other.

The *elementary matrices* come in the same three families, each is the result of performing the corresponding row operation to the identity matrix:

Type I: E_{ij} is the identity matrix with rows i, j swapped;

Type II: $E_i^{(\lambda)}$ is the identity with the i^{th} diagonal entry replaced by $\lambda \neq 0$;

Type III: $E_{ij}^{(\lambda)}$ is the identity matrix with an additional λ in the ij^{th} entry.

Example 3.2. In $M_2(\mathbb{R})$ the elementary matrices are as follows:

$$E_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad E_1^{(\lambda)} = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}, \quad E_2^{(\lambda)} = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}, \quad E_{12}^{(\lambda)} = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}, \quad E_{21}^{(\lambda)} = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}$$

By subtracting three times the first row from the second, we see that the following are row equivalent:

$$\begin{pmatrix} 1 & 4 \\ -2 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 4 \\ -5 & -9 \end{pmatrix}$$

The crucial observation, stated in general below, is that this transformation is the result of multiplying by the corresponding elementary matrix:

$$E_{21}^{(-3)} \begin{pmatrix} 1 & 4 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ -5 & -9 \end{pmatrix}$$

Theorem 3.3. *Let* T *be an elementary row operation acting on* $m \times n$ *matrices.*

1. T *is an isomorphism of* $M_{m\times n}(\mathbb{F})$ *with itself. Its inverse is an operation of the same type.*
2. $\mathrm{T}(A) = EA$ *where* E *is the elementary matrix* $\mathrm{T}(I_m)$ *obtained by applying* T *to the identity.*

In particular, the inverses of the three types of elementary matrix are

$$E_{ij}^{-1} = E_{ij}, \qquad \left(E_i^{(\lambda)}\right)^{-1} = E_i^{(\lambda^{-1})}, \qquad \left(E_{ij}^{(\lambda)}\right)^{-1} = E_{ij}^{(-\lambda)}$$

Proof. Note first that row operations never mix columns and neither does matrix multiplication: if $\mathbf{a}_1, \ldots, \mathbf{a}_n$ are the columns of A, then

$$\mathrm{T}(A) = \big(\mathrm{T}(\mathbf{a}_1) \cdots \mathrm{T}(\mathbf{a}_n)\big) \quad \text{and} \quad EA = \big(E\mathbf{a}_1 \cdots E\mathbf{a}_n\big)$$

It is therefore enough to prove the result when $n = 1$.

1. Suppose that T is of type III, adding to row i a multiple λ of row j. Let v_r, w_r denote the r^{th} entries of $\mathbf{v}, \mathbf{w} \in \mathbb{F}^m$ and let k be a scalar. The r^{th} entry of $\mathrm{T}(k\mathbf{v} + \mathbf{w})$ is plainly

$$\begin{cases} (kv_i + w_i) + \lambda(kv_j + w_j) = k(v_i + \lambda v_j) + (w_i + \lambda w_j) & \text{if } r = i \\ kv_r + w_r & \text{if } r \neq i \end{cases}$$

   which is certainly the r^{th} entry of $k\mathrm{T}(\mathbf{v}) + \mathrm{T}(\mathbf{w})$: thus $\mathrm{T} : \mathbb{F}^m \to \mathbb{F}^m$ is linear. Its inverse is plainly computed by *subtracting* from the i^{th} row λ times the j^{th}: an elementary operation of the same type. Operations of types I and II are similar.

2. By part 1, $\mathrm{T} = \mathrm{L}_E \in \mathcal{L}(\mathbb{F}^m)$ for some invertible matrix $E \in M_m(\mathbb{F})$. To compute E, simply apply T to the standard basis $\epsilon = \{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$:

$$E = \big([\mathrm{T}(\mathbf{e}_1)]_\epsilon \cdots [\mathrm{T}(\mathbf{e}_m)]_\epsilon\big) = \big(\mathrm{T}(\mathbf{e}_1) \cdots \mathrm{T}(\mathbf{e}_m)\big) = \mathrm{T}(\mathbf{e}_1 \cdots \mathbf{e}_n) = \mathrm{T}(I_m)$$

■

Column Operations Applying the above approach to columns yields the *elementary column operations.* Theorem 3.3 holds for column operations provided you multiply by matrices *on the right* $\mathrm{T}(A) = AE$ and replace 'row' with 'column.'

Example 3.4. Let $E = E_{21}^{(3)} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and compute:

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a + 3b & b & c \\ d + 3e & e & f \end{pmatrix}$$

E can be produced from the identity matrix by adding three times the second column to the first, precisely the effect it has as a column operation when multiplying on the right.

Exercises 3.1 1. Let $A = \begin{pmatrix} 2 & 4 \\ 1 & 7 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 7 \\ 0 & 1 \end{pmatrix}$.

   (a) Find a sequence of elementary matrices E^I, E^{II}, E^{III}, of the types indicated, so that

   $$B = E^{II} E^I E^{III} A$$

   (b) Hence find a matrix C such that $B = CA$. Is C the *only* matrix satisfying this equation?

   (c) Find another sequence of elementary matrices such that $B = E_k \cdots E_1 A$.

2. Let A be an $m \times n$ matrix. Prove that if B can be obtained from A by an elementary row operation, then B^T can be obtained from A^T by the corresponding elementary *column* operation. (*This essentially proves Theorem 3.3 for column operations.*)

3. For the matrices A, B in question 1, find a sequence of elementary matrices of any length/type such that $B = AE_1 \cdots E_k$.

3.2 The Rank of a Matrix and Matrix Inverses

Our goal is to use the row equivalence of matrices to provide systematic methods for computing ranks and inverses of linear maps. First we translate the notions of rank and nullity to matrices.

Definition 3.5. The *rank* and *nullity* of a matrix $A \in M_{m\times n}(\mathbb{F})$ are the rank/nullity of the linear map $\mathrm{L}_A : \mathbb{F}^n \to \mathbb{F}^m$ given by left-multiplication by A.

Our previous injectivity and surjectivity conditions immediately translate to this new language.

Lemma 3.6. *Let $A \in M_{m\times n}(\mathbb{F})$ and $\mathrm{L}_A : \mathbb{F}^n \to \mathbb{F}^m$ be the corresponding linear map.*

1. L_A *is injective* $\iff \operatorname{rank} A = n \iff \operatorname{null} A = 0$
2. L_A *is surjective* $\iff \operatorname{rank} A = m \iff \operatorname{null} A = n - m$
3. *(When $m = n$)* L_A *is an isomorphism* $\iff$ *A is invertible* $\iff \operatorname{rank} A = n \iff \operatorname{null} A = 0$

We now come to the crucial observations that permit easy calculations of ranks and inverses.

Theorem 3.7. *Let $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n) \in M_{m\times n}(\mathbb{F})$ have columns $\mathbf{a}_j \in \mathbb{F}^n$.*

1. *The* column space $\operatorname{Span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ *of A has dimension* $\operatorname{rank} A$.
2. $\operatorname{rank} A$ *is invariant under multiplication by invertible matrices: if P and Q are invertible, then*

$$\operatorname{rank} PA = \operatorname{rank} AQ = \operatorname{rank} A$$

Proof.

1. For any vector $\mathbf{v} \in \mathbb{F}^n$, write $\mathbf{v} = \lambda_1\mathbf{e}_1 + \cdots + \lambda_n\mathbf{e}_n$ with respect to the standard basis and observe that

$$\lambda_1\mathbf{a}_1 + \cdots + \lambda_n\mathbf{a}_n = A\mathbf{v} = \mathrm{L}_A(\mathbf{v})$$

It is clear from this that the column space and $\mathcal{R}(L_A)$ are identical, whence $\operatorname{rank} A$ is the dimension of the column space.

2. We work with the range of the linear map $\mathrm{L}_{PA} = \mathrm{L}_P\mathrm{L}_A (= \mathrm{L}_P \circ \mathrm{L}_A)$:

$$\mathcal{R}(\mathrm{L}_{PA}) = \mathcal{R}(\mathrm{L}_P\mathrm{L}_A) = \{PA\mathbf{v} : \mathbf{v} \in \mathbb{F}^n\} = \mathrm{L}_P(\mathcal{R}(\mathrm{L}_A))$$

Since P is invertible, $\mathrm{L}_P : \mathcal{R}(\mathrm{L}_A) \to \mathrm{L}_P(\mathcal{R}(\mathrm{L}_A))$ is an isomorphism, whence

$$\dim \mathcal{R}(\mathrm{L}_A) = \dim \mathrm{L}_P(\mathcal{R}(\mathrm{L}_A)) = \dim \mathcal{R}(\mathrm{L}_{PA}) \implies \operatorname{rank} A = \operatorname{rank} PA$$

The result for AQ is even easier: we leave it to an exercise. ∎

By virtue of the theorem, we'll denote the column space and nullspace of A by $\mathcal{R}(A)$ and $\mathcal{N}(A)$ respectively rather than the lengthier $\mathcal{R}(\mathrm{L}_A)$ and $\mathcal{N}(\mathrm{L}_A)$.

Computing the Rank of a Matrix Recall that elementary row/column operations act via multiplication by invertible matrices: thus

Elementary row/column operations are rank-preserving

Examples 3.8. 1. Recall Example 3.2, where we saw the row equivalence of $\left(\begin{smallmatrix}1&4\\-2&3\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1&4\\-5&-9\end{smallmatrix}\right)$. Since the columns of these are linearly independent, the column spaces of both are $\mathbb{R}^2$ and both matrices plainly have rank 2. Indeed we can perform a sequence of row operations that make the rank even more obvious:

$$\begin{pmatrix}1&4\\-2&3\end{pmatrix}\xrightarrow{E_{21}^{(2)}}\begin{pmatrix}1&4\\0&11\end{pmatrix}\xrightarrow{E_2^{(\frac{1}{11})}}\begin{pmatrix}1&4\\0&1\end{pmatrix}\xrightarrow{E_{12}^{(-4)}}\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

Since all matrices have the same rank, the original clearly has rank 2.

2. Since 2×2 matrices are small, the row operation approach wasn't required. For a larger matrix however, it can be invaluable. For instance:

$$A=\begin{pmatrix}1&1&0&3\\2&0&1&1\\0&1&2&0\\0&2&5&-1\\2&1&1&3\end{pmatrix}\xrightarrow{E_{21}^{(-2)}E_{51}^{(-2)}}\begin{pmatrix}1&1&0&3\\0&-2&1&-5\\0&1&2&0\\0&2&5&-1\\0&-1&1&-3\end{pmatrix}\xrightarrow{E_{13}^{(-1)}E_{23}^{(2)}E_{43}^{(-2)}E_{53}^{(1)}}\begin{pmatrix}1&0&-2&3\\0&0&5&-5\\0&1&2&0\\0&0&1&-1\\0&0&3&-3\end{pmatrix}$$

$$\xrightarrow{E_{14}^{(2)}E_{24}^{(-5)}E_{34}^{(-2)}E_{54}^{(-3)}}\begin{pmatrix}1&0&0&1\\0&0&0&0\\0&1&0&2\\0&0&1&-1\\0&0&0&0\end{pmatrix}\xrightarrow{E_{34}E_{23}}\begin{pmatrix}1&0&0&1\\0&1&0&2\\0&0&1&-1\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

Since the fourth column is a linear combination of the first three (linearly independent) columns, we conclude that $\operatorname{rank} A = 3$. Alternatively, we could repeat using with column operations:

$$\begin{pmatrix}1&1&0&3\\2&0&1&1\\0&1&2&0\\0&2&5&-1\\2&1&1&3\end{pmatrix}\longrightarrow\cdots\longrightarrow\begin{pmatrix}1&0&0&0\\0&0&1&0\\-4&5&-3&0\\-10&12&-7&0\\0&1&0&0\end{pmatrix}$$

The first three columns are linearly independent, so again the rank is three. If we used a mixture of row and column operations, we could eventually transform A into the rank 3 matrix

$$\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$$

Row or Column Operations: which to use? If your only purpose is to compute ranks, mixing row and column operations is fine. If you want something else, you may wish to stick to only one type: here's why.

Row Operations preserve the span of the *rows* of a matrix (the *row space*). This is important when matrices represent linear systems of equations. For example, below we transform a system of equations and the corresponding augmented matrix using row operations:

$$\begin{array}{lll}
\begin{cases} x+3y=1 \\ 2x-y=16 \end{cases} & \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 16 \end{pmatrix} & \left(\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & -1 & 16 \end{array}\right) \\
\begin{cases} x+3y=1 \\ -7y=14 \end{cases} & \begin{pmatrix} 1 & 3 \\ 0 & -7 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 14 \end{pmatrix} & \left(\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & -7 & 14 \end{array}\right) \\
\begin{cases} x+3y=1 \\ y=2 \end{cases} & \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} & \left(\begin{array}{cc|c} 1 & 3 & -1 \\ 0 & 1 & 2 \end{array}\right) \\
\begin{cases} x=-1 \\ y=2 \end{cases} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} & \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 2 \end{array}\right)
\end{array}$$

This familiar method relies on the fact that row-equivalent linear systems have identical solutions. When viewed as a matrix system (middle column) it should be clear that multiplication by elementary matrices must occur *on the left.*

Column Operations preserve the column space of a matrix. For instance, the above example shows that a simple basis of the column space of A is given by

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ -4 \\ -10 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 5 \\ 12 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -3 \\ -7 \\ 0 \end{pmatrix} \right\}$$

Row operations will change the column space and vice versa. If knowing these is important to you, stick to one type of operation!

The example generalizes:

Theorem 3.9. *A matrix $A \in M_{m\times n}(\mathbb{F})$ has* rank $A = r$ *if and only if there exists a finite sequence of row and column operations transforming A to the matrix*

$$D = \begin{pmatrix} I_r & O_{r\times(n-r)} \\ O_{(m-r)\times r} & O_{(m-r)\times(n-r)} \end{pmatrix}$$

Here I_r is the $r \times r$ identity, with the remaining pieces being zero matrices of the given dimensions.

Otherwise said, there exist elementary $m \times m$ matrices $R_1, \ldots, R_k$ and elementary $n \times n$ matrices $C_1, \ldots, C_\ell$ such that

$$R_k \cdots R_1 A C_1 \cdots C_\ell = D$$

The proof is too long to give in full, but can be proved by tedious induction on the number of rows of A. We are far more interested in some corollaries, particularly involving the maximal rank case, that is when A is invertible.

Computing the inverse of a matrix Everything follows from a simple corollary.

Corollary 3.10. *Every invertible matrix A is a product of elementary matrices.*

In light of the Corollary, the last line of Theorem 3.9 can be rewritten so say that

$$\operatorname{rank} A = r \iff \exists\text{invertible } P, Q \text{ such that } PAQ = D$$

Proof. If $A \in M_n(\mathbb{F})$ is invertible, then $\operatorname{rank} A = n$ whence $D = I_n$. It follows that there exist products P, Q of elementary matrices such that

$$PAQ = I_n \implies A = P^{-1}Q^{-1}$$

By Theorem 3.3, the right hand side and thus A is a product of elementary matrices. ∎

The proof yields a systematic method for calculating inverses:

- The identity matrix $I_n = QPA$ is the result of applying a sequence of row operations to A.
- $A^{-1} = QP = QPI_n$ is the result of applying the *same sequence* to I_n.
- Since row operations never mix up columns, we can find A^{-1} by applying row operations to the *augmented matrix* $(A \mid I)$ until the left side is the identity: the right side will then be A^{-1}, i.e.

$$\boxed{(A \mid I) \text{ is row equivalent to } (I \mid A^{-1})}$$

Example 3.11. We compute the inverse of $A = \left(\begin{smallmatrix}1&2&3\\0&2&0\\3&0&1\end{smallmatrix}\right)$ by applying row operations to $(A \mid I_3)$:

$$\left(\begin{array}{ccc|ccc}1&2&3&1&0&0\\0&2&0&0&1&0\\3&0&1&0&0&1\end{array}\right) \xrightarrow{E_{12}^{(-1)}} \left(\begin{array}{ccc|ccc}1&0&3&1&-1&0\\0&2&0&0&1&0\\3&0&1&0&0&1\end{array}\right) \xrightarrow{E_{2}^{(\frac{1}{2})}} \left(\begin{array}{ccc|ccc}1&0&3&1&-1&0\\0&1&0&0&\frac{1}{2}&0\\3&0&1&0&0&1\end{array}\right)$$

$$\xrightarrow{E_{31}^{(-3)}} \left(\begin{array}{ccc|ccc}1&0&3&1&-1&0\\0&1&0&0&\frac{1}{2}&0\\0&0&-8&-3&3&1\end{array}\right) \xrightarrow{E_{3}^{(-\frac{1}{8})}} \left(\begin{array}{ccc|ccc}1&0&3&1&-1&0\\0&1&0&0&\frac{1}{2}&0\\0&0&1&\frac{3}{8}&-\frac{3}{8}&-\frac{1}{8}\end{array}\right)$$

$$\xrightarrow{E_{13}^{(-3)}} \left(\begin{array}{ccc|ccc}1&0&0&-\frac{1}{8}&\frac{1}{8}&\frac{3}{8}\\0&1&0&0&\frac{1}{2}&0\\0&0&1&\frac{3}{8}&-\frac{3}{8}&-\frac{1}{8}\end{array}\right) \implies A^{-1} = \frac{1}{8}\begin{pmatrix}-1&1&3\\0&4&0\\3&-3&-1\end{pmatrix}$$

It can easily be checked by multiplication that we have found the correct inverse matrix: the above indeed shows how to write A as a product of elementary matrices

$$I_3 = E_{13}^{(-3)}E_{3}^{(-\frac{1}{8})}E_{31}^{(-3)}E_{2}^{(\frac{1}{2})}E_{12}^{(-1)}A$$

$$\implies A = E_{12}^{(1)}E_{2}^{(2)}E_{31}^{(3)}E_{3}^{(-8)}E_{13}^{(3)} = \begin{pmatrix}1&1&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&2&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\3&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&-8\end{pmatrix}\begin{pmatrix}1&0&3\\0&1&0\\0&0&1\end{pmatrix}$$

It is also acceptable, though non-standard, to perform column operations on the augmented matrix $\left(\frac{A}{I_n}\right)$: just remember never to mix the two types of operation when computing the inverse!

Corollary 3.12. *For any matrix A we have* $\operatorname{rank} A = \operatorname{rank} A^T$.

Proof. $D = PAQ \implies D^T = Q^T A^T P^T$. Since Q^T and P^T are invertible, we immediately see that $\operatorname{rank} A^T = \operatorname{rank} D^T$. But $\operatorname{rank} D^T$ clearly equals $r = \operatorname{rank} D = \operatorname{rank} A$. ∎

In particular, the dimension of the *row space* (span of the rows of A) is also $\operatorname{rank} A$.

Maximum and Minimum Ranks of Compositions

As a final application we consider the rank of a composition in terms of its factors.

Example 3.13. If $\operatorname{rank} A = 3$ and $\operatorname{rank} B = 2$, we can appeal to block matrices such as in Theorem 3.9 to consider possible ranks of the product AB:

- $A = \begin{pmatrix} 1&0&0\\0&1&0\\0&0&1 \end{pmatrix}$, $B = \begin{pmatrix} 1&0\\0&1\\0&0 \end{pmatrix}$: $AB = B \implies \operatorname{rank} AB = 2$
- $A = \begin{pmatrix} 1&0&0&0\\0&1&0&0\\0&0&1&0 \end{pmatrix}$, $B = \begin{pmatrix} 0&0&0\\0&0&0\\1&0&0\\0&1&0 \end{pmatrix}$: $AB = \begin{pmatrix} 0&0&0\\0&0&0\\1&0&0 \end{pmatrix} \implies \operatorname{rank} AB = 1$
- $A = \begin{pmatrix} 1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0 \end{pmatrix}$, $B = \begin{pmatrix} 0&0&0\\0&0&0\\0&0&0\\1&0&0\\0&1&0 \end{pmatrix}$: $AB = \begin{pmatrix} 0&0&0\\0&0&0\\0&0&0 \end{pmatrix} \implies \operatorname{rank} AB = 0$

As the next result shows, these are essentially all the possibilities.

Theorem 3.14. *Let* $\mathrm{S} : U \to V$ *and* $\mathrm{T} : V \to W$ *be linear maps, then*[a]

$$\operatorname{rank} \mathrm{S} - \operatorname{null} \mathrm{T} \le \operatorname{rank} \mathrm{TS} \le \min\big(\operatorname{rank} \mathrm{T}, \operatorname{rank} \mathrm{S}\big)$$

$$\max\big(\operatorname{null} \mathrm{S}, \dim U - \operatorname{rank} \mathrm{T}\big) \le \operatorname{null} \mathrm{TS} \le \operatorname{null} \mathrm{T} + \operatorname{null} \mathrm{S}$$

The same relationships hold for matrices A, B, provided the product is AB is defined.

[a]The upper bounds are easier to remember due to their symmetry: use the rank–nullity theorem to recover the lower bounds instead of memorizing them!

Proof. If $\mathbf{w} \in \mathcal{R}(\mathrm{TS})$, then $\mathbf{w} = \mathrm{T}(\mathrm{S}(\mathbf{u}))$ for some $\mathbf{u} \in U$, from which $\mathbf{w} \in \mathcal{R}(\mathrm{T})$. We conclude that

$$\mathcal{R}(\mathrm{TS}) \le \mathcal{R}(\mathrm{T}) \implies \operatorname{rank} \mathrm{TS} \le \operatorname{rank} \mathrm{T}$$

If this were a claim about matrices, Corollary 3.12 could deal with the other part of the minimum: $\operatorname{rank} AB = \operatorname{rank}(AB)^T = \cdots$. Instead we consider null spaces and apply the rank-nullity theorem:

$$\mathbf{u} \in \mathcal{N}(\mathrm{S}) \implies \mathrm{S}(\mathbf{u}) = \mathbf{0} \implies \mathrm{TS}(\mathbf{u}) = \mathbf{0} \implies \mathbf{u} \in \mathcal{N}(\mathrm{TS})$$

from which

$$\begin{aligned} \mathcal{N}(\mathrm{S}) \le \mathcal{N}(\mathrm{TS}) &\implies \operatorname{null} \mathrm{S} \le \operatorname{null} \mathrm{TS} \implies \dim U - \operatorname{rank} \mathrm{S} \le \dim U - \operatorname{rank} \mathrm{TS} \\ &\implies \operatorname{rank} \mathrm{TS} \le \operatorname{rank} \mathrm{S} \end{aligned}$$

The remaining inequalities are an exercise. ∎

Example 3.15. Let $A \in M_{6\times 5}(\mathbb{R})$ and $B \in M_{5\times 4}(\mathbb{R})$, and suppose that $\operatorname{rank} A = 4$ and $\operatorname{rank} B = 3$. We find the maximum and minimum possible ranks of the product AB and give examples in each case.

First observe that since $\mathrm{L}_A : \mathbb{R}^5 \to \mathbb{R}^6$, the rank–nullity theorem says that $\operatorname{null} A = 5 - \operatorname{rank} A = 1$. Similarly $\operatorname{null} B = 4 - 3 = 1$.

By the Theorem,

$$2 = \operatorname{rank} B - \operatorname{null} A \le \operatorname{rank} AB \le \min(\operatorname{rank} A, \operatorname{rank} B) = 3$$

It is easy to cook up explicit matrices satisfying $\operatorname{rank} AB = 3$ as in the previous example: for instance

$$A = \begin{pmatrix} I_4 & O_{4\times 1} \\ O_{2\times 4} & O_{2\times 1} \end{pmatrix} = \begin{pmatrix} 1&0&0&0&0 \\ 0&1&0&0&0 \\ 0&0&1&0&0 \\ 0&0&0&1&0 \\ 0&0&0&0&0 \\ 0&0&0&0&0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} \implies AB = \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}$$

The idea for maximum ranks is to have the identity submatrices inside A and B overlap as much as possible.

By trying to make the identities overlap as little as possible—essentially squeezing as much of the range of A into the nullspace of A—we should can also create a minimal rank example: for instance, with the same A as above,

$$B = \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \end{pmatrix} \implies AB = \begin{pmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}$$

Exercises 3.2 1. For each of the following matrices, compute the rank and the inverse if it exists:

(a) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & -2 & 4 \\ 1 & 1 & -1 \\ 2 & 4 & 5 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & -1 & 2 \\ 2 & 0 & 1 & 0 \\ 0 & -1 & 1 & -3 \end{pmatrix}$

2. For each of the following linear transforms T, find the matrix of the linear map with respect to the standard bases, determine whether T is invertible, and compute T^{-1}, if it exists.
   (a) $\mathrm{T} : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ defined by $\mathrm{T}(f)(x) = (x+1)f'(x)$
   (b) $\mathrm{T} : \mathbb{R}^3 \to P_2(\mathbb{R})$ defined by $\mathrm{T}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = (a+b+c) + (a-b+c)x + ax^2$
   (c) $\mathrm{T} : M_2(\mathbb{R}) \to \mathbb{R}^4$ defined by $\mathrm{T}(A) = \begin{pmatrix} \operatorname{tr} A \\ \operatorname{tr} A^T \\ \operatorname{tr}(EA) \\ \operatorname{tr}(AE) \end{pmatrix}$ where $E = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
3. (a) Find $A \in M_{3\times 4}(\mathbb{R})$ and $B \in M_{4\times 3}(\mathbb{R})$ such that $\operatorname{rank} A = \operatorname{rank} B = 2$ and $\operatorname{rank} AB = 1$.
   (b) Suppose that $A \in M_{4\times 3}(\mathbb{R})$ and $B \in M_{3\times 5}(\mathbb{R})$ have $\operatorname{rank} A = 2$ and $\operatorname{rank} B = 3$. What is $\operatorname{rank} AB$?
4. (a) Let $\mathrm{S} \in \mathcal{L}(U, V)$ and $\mathrm{T} \in \mathcal{L}(V, W)$. Prove that

$$\operatorname{null} \mathrm{TS} \le \operatorname{null} \mathrm{T} + \operatorname{null} \mathrm{S}$$

   Now apply the Rank–Nullity Theorem to finish the proof of Theorem 3.14.

(b) Let $A = \begin{pmatrix} I_a & O \\ O & O \end{pmatrix}$, $B = \begin{pmatrix} I_b & O \\ O & O \end{pmatrix}$ and $C = \begin{pmatrix} O & O \\ O & I_c \end{pmatrix}$ where a, b, c are the ranks of A, B, C respectively, and O indicates the zero matrix of the appropriate size. Suppose that

$$A \in M_{m\times n}(\mathbb{F}), \qquad B, C \in M_{n\times p}(\mathbb{F})$$

Compute AB and AC and check that

$$\operatorname{rank} AB = \min(\operatorname{rank} A, \operatorname{rank} B) \quad \text{and} \quad \operatorname{rank} AC = \max(0, \operatorname{rank} C - \operatorname{null} A)$$

This shows that the maximal and minimal ranks indicated in the Theorem can actually be achieved; the only caveat being that $\operatorname{rank} C - \operatorname{null} A$ *could be negative!*

5. Prove that for any $m \times n$ matrix A, we have $\operatorname{rank} A = 0 \iff A$ is the zero matrix.

6. (a) Prove that matrices $A, B \in M_{m\times n}(\mathbb{F})$ have the same rank if and only if $B = PAQ$ for some invertible P, Q.

   (b) Suppose that $\mathrm{T} \in \mathcal{L}(V, W)$ is a linear map between finite-dimensional vector spaces. Show that $\operatorname{rank}[\mathrm{T}]_\beta^\gamma$ is independent of the choice of bases β and γ.

   (*Of course, this value equals* $\operatorname{rank} \mathrm{T}$ *itself!*)

7. Write the invertible matrix $A = \begin{pmatrix} 3 & 2 \\ 10 & 6 \end{pmatrix}$ as a product of elementary matrices.

   For a challenge, see if you can do this for a general invertible matrix.

8. Let $\mathrm{T} \in \mathcal{L}(V, W)$ be given and suppose that $\mathrm{P} \in \mathcal{L}(W)$ and $\mathrm{Q} \in \mathcal{L}(V)$ are isomorphisms. Prove that

$$\operatorname{rank} \mathrm{PT} = \operatorname{rank} \mathrm{TQ} = \operatorname{rank} \mathrm{T}$$

   (*Your argument must work in infinite dimensions and thus without matrices*)