1. If a linear transformation \( L : V \rightarrow W \) is invertible, then \( L \) must be bijective, so \( \text{nullity } L = 0 \) and \( \text{rank } L = \dim W \). But if \( L \) is a bijective \( LT \), then it is an isomorphism, so \( V \cong W \) and thus \( \dim V = \dim W \).

In summary, \( \text{rank } L = \dim V = \dim W \), and \( \text{nullity } L = 0 \).

2. \( L : V \rightarrow W \) and \( K : W \rightarrow U \): linear transformations

(a) claim: \( \text{im}(K \circ L) \subseteq \text{im} K \rightarrow \text{i.e., } \forall \mathbf{v} \in \text{im}(K \circ L) \implies \mathbf{v} \in \text{im} K \)

Proof: Suppose \( \mathbf{v} \in \text{im}(K \circ L) \). 
- Recall that \( \text{im}(K \circ L) = \{ (K \circ L)(\mathbf{v}) : \forall \mathbf{v} \in V \} \).
- Then \( \mathbf{v} = (K \circ L)(\mathbf{v}) \), where \( \mathbf{v} \in V \).
- So by definition, \( \mathbf{v} \in \text{im} K \).
- Therefore, \( \forall \mathbf{v} \in \text{im}(K \circ L) \implies \mathbf{v} \in \text{im} K \).

(b) claim: \( \text{rank } (K \circ L) \leq \text{rank } L \)

Proof: Let \( \{ \mathbf{v}_1, \ldots, \mathbf{v}_n \} \) be a basis for \( V \); \( \{ K(\mathbf{v}_1), \ldots, K(\mathbf{v}_n) \} \) are a basis for \( \text{ker } L \) and \( \mathbf{v} = \text{rank } L \), as in the proof of the rank-nullity theorem.

Then since this collection spans \( V \), the collection \( \{ (K \circ L)(\mathbf{v}_1), \ldots, (K \circ L)(\mathbf{v}_n), (K \circ L)(\mathbf{v}_1), \ldots, (K \circ L)(\mathbf{v}_n) \} \)
spans \( \text{im } (K \circ L) \).
- But \( \forall i, (K \circ L)(\mathbf{v}_i) = K(L(\mathbf{v}_i)) \).
- Therefore, \( \forall \mathbf{v}_i \in \text{ker } L \implies (K \circ L)(\mathbf{v}_i) = 0 \).

So \( \{ (K \circ L)(\mathbf{v}_1), \ldots, (K \circ L)(\mathbf{v}_n) \} \) spans \( \text{im } (K \circ L) \),
and thus can be reduced to a basis for \( \text{im } (K \circ L) \); thus, \( \dim \{ \text{im } (K \circ L) \} = \mathbf{v} = \text{rank } L \).
- So by definition, \( \text{rank } (K \circ L) \leq \text{rank } L \).

(c) claim: \( \text{ker } L = \text{ker } (K \circ L) \rightarrow \text{i.e., } \forall \mathbf{v} \in \text{ker } L \implies \mathbf{v} \in \text{ker } (K \circ L) \)

Proof: Let \( \mathbf{v} \in \text{ker } L \); then by definition, \( L(\mathbf{v}) = 0 \).
- Therefore, \( \forall \mathbf{v} \in \text{ker } L \implies (K \circ L)(\mathbf{v}) = 0 \).

Thus, \( \{ (K \circ L)(\mathbf{v}) = K(L(\mathbf{v})) \} \) which implies \( (K \circ L)(\mathbf{v}) = 0 \).
- Therefore, \( \forall \mathbf{v} \in \text{ker } L \implies \mathbf{v} \in \text{ker } (K \circ L) \).

Corollary: \( \text{nullity } (K \circ L) \geq \text{nullity } L \)

Proof: Since \( \text{ker } L \) is a subspace of \( \text{ker } (K \circ L) \), \( \dim \{ \text{ker } L \} \leq \dim \{ \text{ker } (K \circ L) \} \).
- So by definition, \( \text{nullity } L \leq \text{nullity } (K \circ L) \).

Corollary: \( \text{nullity } (K \circ L) \geq \text{nullity } L \)

Proof: Since \( \text{ker } L \) is a subspace of \( \text{ker } (K \circ L) \), \( \dim \{ \text{ker } L \} \leq \dim \{ \text{ker } (K \circ L) \} \).
- So by definition, \( \text{nullity } L \leq \text{nullity } (K \circ L) \).
3. We think of an n-dimensional vector space as having "size" n (formally, this is justified by counting the number of vectors in any basis for it).

We've seen in our proof of the Rank + Nullity Theorem that any linear transformation from one such space to another kills its nullity's worth of these and generates an image of size given by its rank.

E.g., for \( L: \mathbb{R}^5 \rightarrow \mathbb{R}^4 \) of rank 3 (\( \text{nullity } 2 \)),

\[ \text{5-dim' domain} \xrightarrow{L} \text{4-dim' codomain} \]

Kill \((\text{nullity } 2)\) slots' worth of input.

Push the rest of the slots' worth into the codomain.

Following the "size" of the subspaces as they're pushed through gives us the rank of a composition; counting how many are killed gives us the nullity.

\* RCN: A L.T. will kill its nullity's worth of slots!

E.g., \( K \circ L \), where \( L \) is as above.

And \( K: \mathbb{R}^4 \rightarrow \mathbb{R}^3 \) has rank 2:

\[ \text{5-dim' subspace in; 2 slots killed} \]

\[ \text{3 survive; 2 slots killed by } K \text{ -- could kill one or two dimensions of the subspace} \]

Not effect: \( K \circ L \) must have rank 2 or 1.

(Keeping track of the extreme cases shows you the possible range of dimensions)

4. (a) \( \text{Just use } \text{rank} + \text{nullity} = \text{dim of domain} \)

- \( \text{nullity } L = 4 - 3 = 1 \)
- \( \text{nullity } K = 3 - 3 = 0 \)
- \( \text{nullity } M = 5 - 1 = 4 \)
- \( \text{nullity } N = 4 - 4 = 0 \)
- \( \text{nullity } R = 3 - 3 = 0 \)
- \( \text{nullity } J = 5 - 4 = 1 \)

(b) \( L \) is not injective (nullity \( \neq 0 \)), but is surjective (rank \( = 3 \))

- \( K \) is injective (nullity \( = 0 \)), but not surjective (rank \( \neq 3 \))

- \( M \) is not injective (nullity \( \neq 0 \)) and not surjective (rank \( \neq 3 \))

- \( N \) is bijective (nullity \( = 0 \), rank \( = 4 \))

- \( R \) is bijective (nullity \( = 0 \), rank \( = 3 \))

- \( J \) is not injective (nullity \( \neq 0 \)) and not surjective (rank \( \neq 4 \))

(c) (i) \[ L \xrightarrow{\text{nullity } 2} K \xrightarrow{\text{nullity } 0} \]

\( K \circ L \) must have rank 3 (\( \text{nullity } 1 \)).

\( K \circ L \) can be neither injective nor surjective.

(c) (ii) \[ K \xrightarrow{\text{nullity } 0} M \xrightarrow{\text{nullity } +} \]

\( M \circ K \) either has rank 1 (\( \text{nullity } 2 \)) or rank 0 (\( \text{nullity } 3 \)).

\( M \circ K \) can be neither injective nor surjective.
(iii) \[ M \xrightarrow{\text{nullity 4}} \text{MAX} \xrightarrow{\sim} \text{MIN} \] \[ L \xrightarrow{\text{nullity 1}} \text{MIN} \xrightarrow{\sim} \text{MAX} \] \[ \therefore \text{can be neither injective nor surjective} \]

(iv) \[ N \xrightarrow{\text{nullity 0}} \text{MAX} \xrightarrow{\sim} \text{MIN} \] \[ R \xrightarrow{\text{nullity 0}} \text{MIN} \xrightarrow{\sim} \text{MAX} \] \[ \therefore \text{can be surjective, but not injective} \]

(v) \[ J \xrightarrow{\text{nullity 1}} \text{MAX} \xrightarrow{\sim} \text{MIN} \] \[ M \xrightarrow{\text{nullity 4}} \text{MIN} \xrightarrow{\sim} \text{MAX} \] \[ \therefore \text{can be neither injective nor surjective} \]

(vi) \[ N \xrightarrow{\text{nullity 0}} \text{MAX} \xrightarrow{\sim} \text{MIN} \] \[ N \xrightarrow{\text{nullity 0}} \text{MIN} \xrightarrow{\sim} \text{MAX} \] \[ \therefore \text{must be bijective} \]

(vii) \[ R \xrightarrow{\text{nullity 0}} \text{MAX} \xrightarrow{\sim} \text{MIN} \] \[ K \xrightarrow{\text{nullity 0}} \text{MIN} \xrightarrow{\sim} \text{MAX} \] \[ \therefore \text{can't be surjective} \]

(viii) \[ J \xrightarrow{\text{nullity 1}} \text{MAX} \xrightarrow{\sim} \text{MIN} \] \[ J \xrightarrow{\text{nullity 1}} \text{MIN} \xrightarrow{\sim} \text{MAX} \] \[ \therefore \text{can't be injective, can't be surjective} \]

(ix) \[ \text{Pickling up where we left off at J} \xrightarrow{\text{nullity 1}} \text{MAX} \xrightarrow{\sim} \text{MIN} \] \[ J \xrightarrow{\text{nullity 1}} \text{MIN} \xrightarrow{\sim} \text{MAX} \] \[ \therefore \text{can't be injective, can't be surjective} \]
5. A: \(m \times n\) matrix; B: \(n \times p\) matrix
(a) As a L.T., \(A: \mathbb{R}^m \rightarrow \mathbb{R}^m\) and \(B: \mathbb{R}^p \rightarrow \mathbb{R}^m\).
(b) The matrix \(AB\) simply means the matrix for the corresponding composition \(A \circ B: \mathbb{R}^p \rightarrow \mathbb{R}^m\).

Thus, \(AB\) is an \(m \times p\) matrix, and our rule for computing it simply comes from analyzing the composition \(A \circ B\) as a linear transformation.
(c) The \(j^{th}\) column of \(AB\)
   \[ AB \vec{e}_j = A (B \vec{e}_j) = \text{column 3 of } AB \]
   where \(AB\) is a composition
   = linear combination of columns of \(A\) with scalars given by the \(j^{th}\) col. of \(B\).

E.g.,
\[
\begin{bmatrix}
2 & 1 \\
1 & 4 \\
3 & 0
\end{bmatrix}
\begin{bmatrix}
1 & 1 & 0 \\
1 & 0 & 1 \\
2 & 3 & 1
\end{bmatrix} =
\begin{bmatrix}
4 & 15 & -3 \\
9 & 11 & -8 \\
3 & 3 & 12
\end{bmatrix}
\]

6. The identity matrix of size \(n\) is \(I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}\),

which corresponds to the standard basis for \(\mathbb{R}^n\).

This matrix gives us the identity L.T.: \(I_n : \mathbb{R}^n \rightarrow \mathbb{R}^n\), \(I_n x = x\),

so:
- \(I_n x = x\) \(\forall x \in \mathbb{R}^n\)
- \(A I_n = A\) for any \(m \times n\) matrix \(A\)
- \(I_n B = B\) for any \(n \times p\) matrix \(B\).

7. (a) An inverse for a matrix \(A\) is a matrix \(A^{-1}\) for which \(A^{-1} A = I = A A^{-1}\) (where \(I\) is the identity matrix of the appropriate size).

(i.e., \(A\) is invertible if it has an inverse.
(b) In order for \(A\) to be invertible, it must be square, with "full rank" and zero nullity.

(i.e., rank as big as possible — here, this means equal to the # of rows (or columns) of \(A\).

(c) Using the fact that \(A A^{-1} = I\), we can find the \(j^{th}\) column of \(A^{-1}\) as follows:

\[ A A^{-1} \vec{e}_j = \vec{e}_j = \vec{e}_j, \]

so \(A (A^{-1} \vec{e}_j) = \vec{e}_j\).

\(i.e., A (\text{column } j \text{ of } A^{-1}) = \vec{e}_j\),

so the \(j^{th}\) column of \(A^{-1}\) is simply the solution of the system \([A | \vec{e}_j]\).

We can thus find all columns of \(A^{-1}\) at once by reducing the multiply-augmented system \([A | I]\):

\[
[A | I] \rightarrow [I | A^{-1}]
\]

Analyze the result! The columns here are precisely those of \(A^{-1}\!\).

(b) \(\rightarrow [I | A^{-1}]\)

Since \(A\) is square and has full rank, the left side must reduce to this!

If we try this for a non-invertible matrix \(A\), the left side won't come out to \(I\), and the system will be inconsistent in some columns.

(d) Claim: If \(AA^{-1} = I\) and \(BA = I\), then \(B = A^{-1}\).

Proof: Suppose that \(AA^{-1} = I\) and \(BA = I\).

Then right-composing \(BA = I\) with \(A^{-1}\), we have \(BAA^{-1} = IA^{-1}\).

So \(BI = A^{-1}\), by the hypothesis that \(AA^{-1} = I\).

\(\therefore B = A^{-1}\) by properties of \(I\).
8. If we have an inverse $A^{-1}$ for $A$, then

$A\mathbf{x} = \mathbf{b} \implies A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b}$ (comparing with $A^{-1}$ on the left)

$\implies (A^{-1}A)\mathbf{x} = A^{-1}\mathbf{b}$

$\implies \mathbf{I}\mathbf{x} = A^{-1}\mathbf{b}$ by definition of inverse

$\implies \mathbf{x} = A^{-1}\mathbf{b}$ by properties of $\mathbf{I}$

— i.e., just let $A^{-1}$ act on $\mathbf{b}$

*However, the initial "if" makes this an extremely specialized approach: it only works for "square" systems, having the same number of relations + unknowns, and it only works if the coefficient matrix is invertible!