

This handout contains a “grand summary” of all we have studied in this course so far. We consider a system of linear equations, written in matrix form:

$$A\mathbf{x} = \mathbf{b}$$

A is a $m \times n$ matrix, \mathbf{x} is a $n \times 1$ vector, and \mathbf{b} is a $m \times 1$ vector.

With this system, we associate a linear transformation $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ in the usual way:

$$L(\mathbf{x}) = A\mathbf{x} \text{ for } \mathbf{x} \text{ in } \mathbb{R}^n.$$

We start, in the next section, by reviewing some key definitions.

1 Essential Concepts

It is convenient, to make it easier to state all the results relationships that come up, to introduce some concepts that use the modern language of set theory. In the following definitions, $L : V \rightarrow W$ is a linear transformation, and V and W are vector spaces.

Definition 1.1. The *range* of L is the set of all vectors \mathbf{b} in \mathbb{R}^m for which there is a vector \mathbf{x} in \mathbb{R}^n such that $L(\mathbf{x}) = \mathbf{b}$. The range of L is a subspace of \mathbb{R}^m , denoted by $\text{range}(L)$. The *rank* of L , denoted by $\text{rank}(L)$ is defined as the dimension of $\text{range}(L)$.

Definition 1.2. The *kernel* of L is the set of all vectors \mathbf{x} in \mathbb{R}^n such that $L(\mathbf{x}) = \mathbf{0}$. The kernel of L is denoted by $\text{kernel}(L)$. The *nullity* of L , denoted by $\text{nullity}(L)$ is the dimension of $\text{kernel}(L)$.

The concepts of range, kernel, rank and nullity have analogous definitions for a matrix A . So, for example $\text{range}(A)$ and $\text{kernel}(A)$ are, respectively, the same as $\text{range}(L)$ and $\text{kernel}(L)$, where L is the corresponding linear transformation¹. Notice that, if A is $m \times n$ then L maps \mathbb{R}^n into \mathbb{R}^m .

It is important to realize that these concepts are just a new way to describe properties of a system of linear equations:

- $\text{range}(L)$ is the set of all vectors \mathbf{b} for which the system $A\mathbf{x} = \mathbf{b}$ has a solution.
- $\text{kernel}(L)$ is the set of all solutions \mathbf{x} of the system $A\mathbf{x} = \mathbf{0}$.

Also notice that we already have methods of computing the range and kernel. We proceed as follows:

¹The attentive reader will notice that there may be an ambiguity here, since we could represent the transformation L in a different bases, getting a different matrix A . It can be shown that the definitions are consistent, no matter what are the bases that are chosen to represent L .

1. R be the reduced row echelon form of A .
2. A basis of $\text{range}(A)$ is given by the columns of A that correspond to pivot columns of R . Consequently, $\text{rank}(A)$ is the number of pivot columns in R .
3. A basis of $\text{kernel}(A)$ is obtained by the usual procedure of finding a set of solutions of $R\mathbf{x} = \mathbf{0}$. It follows that $\text{nullity}(A)$ is the number of free variable columns in R .

This procedure has a very important consequence. The $\text{rank}(A)$ is equal to the number of columns of A corresponding to pivot variables, and $\text{nullity}(A)$ is equal to the number of columns of A corresponding to free variables. Since every column corresponds to either a free variable or a pivot variable we have:

$$\text{rank}(A) + \text{nullity}(A) = \text{Number of columns of } A$$

So, we have the following very important theorem:

Theorem 1.3 (Rank-Nullity Theorem). *Let $L : V \rightarrow W$. Then:*

$$\text{rank}(L) + \text{nullity}(L) = \dim(\text{range}(L)) + \dim(\text{kernel}(L)) = \dim(V)$$

In terms of matrices:

$$\text{rank}(A) + \text{nullity}(A) = \dim(\text{range}(A)) + \dim(\text{kernel}(A)) = \text{Number of columns of } A$$

2 One-to-one and onto

The two most important questions we can ask about a system of linear equations:

$$A\mathbf{x} = \mathbf{b}$$

are the following:

- **Existence.** Does the system have solutions \mathbf{x} for a given \mathbf{b} ?
- **Uniqueness.** Given that the system has solutions, is there only one solution?

These questions can be phrased in terms of linear transformations. Suppose that $L : V \rightarrow W$ is an arbitrary linear transformation, where V and W are vector spaces. (You can think of $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$, but the definitions apply to arbitrary spaces.)

Definition 2.1. We say that the linear transformation L is *onto* if for every vector \mathbf{b} in W , there is a vector \mathbf{x} in V such that $L(\mathbf{x}) = \mathbf{b}$.

Definition 2.2. We say that the linear transformation L is *one-to-one* if, for every \mathbf{b} in W , there is at most one \mathbf{x} in V such that $L(\mathbf{x}) = \mathbf{b}$.

As usual, we can interpret these definitions in terms of systems. Suppose that A is a $m \times n$ matrix, and let L be the corresponding linear transformation. Then:

- L is onto if and only if, for every \mathbf{b} in \mathbb{R}^m , the system $A\mathbf{x} = \mathbf{b}$ has a solution.
- L is one-to-one if and only if, for every \mathbf{b} in \mathbb{R}^m , the system $A\mathbf{x} = \mathbf{b}$ has at most one solutions. That is, if a solution exists, then it is unique.

Now come the first important results of this handout. The notions of onto and one-to-one can be expressed in terms of the subspaces we defined in the previous section. Let's start with a characterization of linear transformations that are onto.

Theorem 2.3. *Let $L : V \rightarrow W$ be a linear transformation. Then, L is onto if and only if any of the following statements is true:*

1. $\text{range}(L) = W$
2. $\text{rank}(L) = \dim(W)$
3. $\text{nullity}(L) = \dim(V) - \dim(W)$

Proof. The first statement is just a restatement of the definition of onto.

To see why the second statement is equivalent to the first, notice that $\text{range}(L) = W$ if and only if $\text{rank}(L) = \dim(\text{range}(L)) = \dim(W)$.

To see that the third statement is equivalent to the second, recall the rank-nullity theorem:

$$\text{rank}(L) + \text{nullity}(L) = \dim(V).$$

So, $\text{rank}(L) = \dim(W)$ if and only if:

$$\dim(W) + \text{nullity}(L) = \dim(V).$$

□

The third statement has an important consequence.

Corollary 2.4. *Let $L : V \rightarrow W$ be a linear transformation. If $\dim(V) < \dim(W)$, then L cannot be onto.*

Proof. By the previous theorem, L is onto if and only if $\text{nullity}(L) = \dim(V) - \dim(W)$. But, if $\dim(V) < \dim(W)$, this would imply $\text{nullity}(L) < 0$, which is impossible. □

Example 2.5. Show that there are vectors $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$ for which the system:

$$\begin{bmatrix} 1 & 2 & 0 \\ -3 & 4 & 2 \\ 0 & 5 & -10 \\ 2 & 2 & 13 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

does not have solutions.

Solution: The linear transformation L associated to A maps \mathbb{R}^3 into \mathbb{R}^4 , since A is 4×3 . Since $\dim(\mathbb{R}^4) > \dim(\mathbb{R}^3)$, L cannot be onto. It follows that there are vectors \mathbf{b} that are not the image $L(\mathbf{x})$ of any \mathbf{x} in \mathbb{R}^3 . For such vectors \mathbf{b} , the system does not have any solution.

We now turn to the characterization of one-to-one linear transformations.

Theorem 2.6. Let $L : V \rightarrow W$ be a linear transformation, where V and W are arbitrary vector spaces. Then, L is one-to-one if and only if any of the following statements are true:

1. $\text{kernel}(L) = \{\mathbf{0}\}$.
2. $\text{nullity}(L) = 0$.
3. $\text{rank}(L) = \dim(V)$.

Proof. Let's start by proving that the first statement is equivalent to L being one-to-one. Since we want to prove a “if and only if” statement, we break it down into two implications.

First, suppose that L is one-to-one. If \mathbf{x} is in $\text{kernel}(L)$, then, $L(\mathbf{x}) = \mathbf{0} = L(\mathbf{0})$, and L being one-to-one implies $\mathbf{x} = \mathbf{0}$. This shows that $\text{kernel}(L) = \{\mathbf{0}\}$.

Now, suppose that $\text{kernel}(L) = \{\mathbf{0}\}$. This is equivalent to the following:

$$L(\mathbf{z}) = \mathbf{0} \text{ implies } \mathbf{z} = \mathbf{0}.$$

To show that L is one-to-one, suppose that $L(\mathbf{x}) = L(\mathbf{y})$. Then, using linearity, we get:

$$L(\mathbf{x} - \mathbf{y}) = L(\mathbf{x}) - L(\mathbf{y}) = \mathbf{0}$$

Using $\mathbf{z} = \mathbf{x} - \mathbf{y}$ in the previous statement, we get $\mathbf{x} - \mathbf{y} = \mathbf{0}$, that is, $\mathbf{x} = \mathbf{y}$. This implies that L is one-to-one.

The proof that the second statement is equivalent to the first is almost immediate. Just notice that $\text{kernel}(L) = \{\mathbf{0}\}$ if and only if $\text{nullity}(L) = \dim(\text{kernel}(L)) = 0$.

To prove that the third statement is equivalent to the second, we use the rank-nullity theorem again:

$$\text{rank}(L) + \text{nullity}(L) = \dim(V).$$

From this, it follows immediately that $\text{nullity}(L) = 0$ if and only if $\text{rank}(L) = \dim(V)$. \square

Example 2.7. Show that the system

$$\begin{bmatrix} 1 & 2 & -2 \\ -3 & 0 & 8 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

has nonzero solutions.

Solution: Let L be the linear transformation associated with the matrix A in the given system. Then, L maps \mathbb{R}^3 to \mathbb{R}^2 , since A is 2×3 . Notice that the solution set of the system is simply $\ker(L)$.

Notice that:

- By the third statement in Theorem ??, L is one-to-one if and only if $\text{rank}(L) = \dim(\mathbb{R}^3) = 3$.
- $\text{rank}(L) = \dim(\text{range}(L))$, and $\text{range}(L)$ is a subspace of \mathbb{R}^2 . So, $\text{rank}(L) \leq 2$.

The two statements above imply that L can't be one-to-one. By the first statement in Theorem ??, this implies that there are nonzero vectors in $\ker(L)$. Any such vector will be a nonzero solution of the given system.

The reasoning above is valid for a general matrix A , and has the following important consequence:

Theorem 2.8. *Suppose that A is a $m \times n$ matrix where $n > m$. Then the homogeneous system $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions.*

Proof. Let $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be the linear transformation associated to A . Then, L is one-to-one if and only if $\text{rank}(L) = n > m$, which is impossible. So, L is not one-to-one, and $\ker(L)$ has nonzero vectors. Since $\ker(L)$ is a subspace, this implies that $\ker(L)$ is an infinite set. Since the solution set of the system is just the kernel of L , this shows that the system has infinitely many solutions. \square

Note: We had already proved this result by analyzing the RREF of A .

3 The Alternatives Theorems

We can now state very general theorems that characterize the structure of the solution sets of linear system. The theorems in this section simply summarize the results from previous sections, so their proofs are omitted.

In the results below, we assume that A is a $m \times n$ matrix.

Theorem 3.1 (Alternatives Theorem for Homogeneous Linear Systems). *Consider the system*

$$A\mathbf{x} = \mathbf{0}$$

Then exactly one of the following alternatives is true:

1. *The system has exactly one solution.*
2. *The system has infinitely many solutions.*

Furthermore, we have the following facts:

- *Alternative 1 is true if and only if $\text{nullity}(L) = 0$ or, equivalently, $\text{rank}(L) = n$.*
- *If $n > m$, Alternative 2 is always true.*

For general systems we have the following.

Theorem 3.2 (Alternatives Theorem for General Linear Systems). *Consider the system*

$$A\mathbf{x} = \mathbf{b}$$

Then, for each vector \mathbf{b} in \mathbb{R}^m , exactly one of the following alternatives takes place:

1. *The system has exactly one solution.*
2. *The system has infinitely many solutions.*
3. *The system has no solutions.*

Furthermore, the following statements are true:

- *There are vectors \mathbf{b} for which Alternative 1 is true if and only if $\text{nullity}(L) = 0$ or, equivalently, $\text{rank}(L) = n$.*
- *If $\text{nullity}(L) > 0$ or, equivalently, $\text{rank}(L) < n$, then there are vectors that satisfy Alternative 2.*
- *If $n > m$, then there are vectors for which Alternative 2 is true and Alternative 1 is not possible.*
- *If $n < m$, then there are vectors for which Alternative 3 is true.*
- *If $n \neq m$, then it is not possible that Alternative 1 is true for all vectors \mathbf{b} .*

4 Square matrices

An important case occurs when the the matrix A is square. Then we can give a more precise characterization of the solution set of the system.

Theorem 4.1 (Alternatives Theorem for Square Matrices). *Let A be a $n \times n$ matrix, and consider the system:*

$$A\mathbf{x} = \mathbf{b}$$

Then, the exactly one of the following alternatives takes place:

- 1. For all vectors \mathbf{b} , the system has exactly one solution.*
- 2. There are vectors \mathbf{b} for which the system has no solution, and there are vectors \mathbf{b} for which the system has infinitely many solutions.*

Furthermore, Alternative 1 is true if and only if any of the following equivalent statements is true:

- The only solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.*
- $\text{nullity}(A) = 0$*
- $\text{rank}(A) = n$*
- $\det(A) \neq 0$*