Linear Algebra: Lecture 3

More on linear transformations. Topics: range and null space, linear equation, linear equations in $\mathbb{R}^n$.

Range and null space.

There are two very important subspaces associated with a linear transformation $T : \mathcal{V} \to \mathcal{W}$.

- The *range* of T, $\mathcal{R}(T) = \{w \in \mathcal{W} : \ w = Tv \text{ for some } v \in \mathcal{V}\}$.
- The *null space* of T, $\mathcal{N}(T) = \{v \in \mathcal{V} : \ Tv = 0\}$.

We say that T is *onto* if $\mathcal{R}(T) = \mathcal{W}$ and that T is *one-to-one* (abbreviated 1-1) if $T(u) = T(v)$ only when $u = v$. We (sometimes) say that T is an *isomorphism* between $\mathcal{V}$ and $\mathcal{W}$ if it is both 1-1 and onto.

Suppose that $T : \mathcal{V} \to \mathcal{W}$ is 1-1 and onto. Then we can define an *inverse* map $T^{-1} : \mathcal{W} \to \mathcal{V}$ as follows: For each $w \in \mathcal{W}$, define $T^{-1}(w)$ to be the unique $v \in \mathcal{V}$ such that $T(v) = w$.

Note that $T^{-1}(T(v)) = v$ for all $v \in \mathcal{V}$ and $T(T^{-1}(w)) = w$ for all $w \in \mathcal{W}$.

It is easy to see that $T^{-1} : \mathcal{W} \to \mathcal{V}$ is linear. Indeed, suppose we have w_1, $w_2 \in \mathcal{W}$ and scalars α_1, α_2. Set $v_1 = T^{-1}(w_1)$ and $v_2 = T^{-1}(w_2)$. We have that $T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2) = \alpha_1 w_1 + \alpha_2 w_2$, and it follows that $T^{-1}(\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 v_1 + \alpha_2 v_2 = \alpha_1 T^{-1}(w_1) + \alpha_2 T^{-1}(w_2)$.

Conversely, suppose that T has an inverse, i.e., a linear map $T^{-1} : \mathcal{W} \to \mathcal{V}$ such that $T^{-1}(T(v)) = v$ for all $v \in \mathcal{V}$ and $T(T^{-1}(w)) = w$ for all $w \in \mathcal{W}$. Then (1) T is 1-1, since $T(u) = T(v) \Rightarrow T^{-1}(T(u)) = T^{-1}(T(v)) \Rightarrow u = v$; (2) T is onto, since, for every $w \in \mathcal{W}$, we have that $T(v) = w$ for $v = T^{-1}(w) \in \mathcal{V}$.

The following summarizes these observations.

PROPOSITION 3.1. *$T : \mathcal{V} \to \mathcal{W}$ is 1-1 and onto if and only if there is a linear map $T^{-1} : \mathcal{W} \to \mathcal{V}$ such that $T^{-1}(T(v)) = v$ for all $v \in \mathcal{V}$ and $T(T^{-1}(w)) = w$ for all $w \in \mathcal{W}$.*

The following provides useful characteristic properties of 1-1 maps.

PROPOSITION 3.2. *The following are equivalent:*

(a) *T is 1-1.*

(b) *$\mathcal{N}(T) = \{0\}$, i.e., $T(v) = 0 \iff v = 0$.*

(c) *T maps linearly independent sets to linearly independent sets, i.e., if $\{v_1, \ldots, v_k\}$ is linearly independent, then so is $\{T(v_1), \ldots, T(v_k)\}$.*

Proof.
(a)⇒(b) Since T is 1-1, $T(v) = 0$ only when $v = 0$, so 0 is the only vector in $\mathcal{N}(T)$.

(b)⇒(c) Suppose that $\mathcal{N}(T) = \{0\}$. Let $\{v_1, \ldots, v_k\}$ be linearly independent in $\mathcal{V}$. If $\alpha_1 T(v_1) + \ldots + \alpha_n T(v_n) = 0$, then $T(\alpha_1 v_1 + \ldots + \alpha_n v_n) = 0$ since T is linear. It follows that $\alpha_1 v_1 + \ldots + \alpha_n v_n \in \mathcal{N}(T)$ and, since $\mathcal{N}(T) = \{0\}$, that $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$. Then, since $\{v_1, \ldots, v_k\}$ is linearly independent, we must have that $\alpha_1 = \ldots = \alpha_n = 0$.

(c)⇒(a) Suppose that T maps linearly independent sets to linearly independent sets. Since a non-zero $v \in \mathcal{V}$ constitutes a linearly independent set, we have in particular that $T(v) \neq 0$ whenever $v \neq 0$, or, equivalently, that $T(v) = 0$ only if $v = 0$. It follows that if $T(u) = T(v)$ for u, $v \in \mathcal{V}$, then $T(u-v) = 0$ only if $u - v = 0$, i.e., $u = v$. $\square$

We can say more if $\mathcal{V}$ and perhaps $\mathcal{W}$ are finite-dimensional.

Proposition 3.3. *Suppose that* $\dim \mathcal{V} = n$. *If* $\{v_1, \ldots, v_n\}$ *is a basis for* $\mathcal{V}$, *then* $\{T(v_1), \ldots, T(v_n)\}$ *is a spanning set for* $\mathcal{R}(T)$.

Proof. If $w \in \mathcal{R}(T)$, then $w = T(v)$ for some $v \in \mathcal{V}$. Writing $v = \sum_{i=1}^{n} \alpha_i v_i$, we have that $w = T(\sum_{i=1}^{n} \alpha_i v_i) = \sum_{i=1}^{n} \alpha_i T(v_i) \in \operatorname{span}\{T(v_1), \ldots, T(v_n)\}$. $\square$

It follows from Proposition 3.3 that $\dim \mathcal{R}(T) \leq n$. Also, if $\dim \mathcal{W} = m$, then we clearly have that $\dim \mathcal{R}(T) \leq m$; moreover, T is *onto* if and only if $\dim \mathcal{R}(T) = m$.

We now work toward Theorem 3.6 below. This is a fundamentally important result for linear transformations on finite-dimensional vector spaces. We will use it in the sequel to gain basic insights into the existence and uniqueness of solutions of linear equations.

We assume throughout the following that $\mathcal{V}$ is finite-dimensional, with $\dim \mathcal{V} = n$, and that T is a linear transformation from $\mathcal{V}$ to $\mathcal{W}$.

Suppose that $\dim \mathcal{N}(T) = k$ for some k, $0 \leq k \leq n$. Let $\{v_1, \ldots, v_k\}$ be a basis for $\mathcal{N}(T)$. By Proposition 2.7, we can expand this to a basis $\{v_1, \ldots, v_k, u_1, \ldots, u_{n-k}\}$ for $\mathcal{V}$. Set $\mathcal{M} = \operatorname{span}\{u_1, \ldots, u_{n-k}\}$. Clearly, $\{u_1, \ldots, u_{n-k}\}$ is linearly independent and so is a basis for $\mathcal{M}$; thus $\dim \mathcal{M} = n - k$.

Note: $\mathcal{M}$ is not unique, since the expansion $\{u_1, \ldots, u_{n-k}\}$ is not unique. However, every such expansion of $\{v_1, \ldots, v_k\}$ will result in a subspace of dimension $n - k$, which is the main thing that matters to us.

Proposition 3.4. *With* $\mathcal{N}(T)$ *and* $\mathcal{M}$ *as above,* $\mathcal{N}(T) \cap \mathcal{M} = \{0\}$.

Proof. Clearly, $\{0\} \subseteq \mathcal{N}(T) \cap \mathcal{M}$. Conversely, if $v \in \mathcal{N}(T) \cap \mathcal{M}$, then, since $\{v_1, \ldots, v_k\}$ is a basis of $\mathcal{N}(T)$, we can write $v = \sum_{i=1}^{k} \alpha_i v_i$ and, since $\{u_1, \ldots, u_{n-k}\}$ is a basis of $\mathcal{M}$, we can also write $v = \sum_{i=1}^{n-k} \beta_i u_i$. Then $\sum_{i=1}^{k} \alpha_i v_i = \sum_{i=1}^{n-k} \beta_i u_i$, which yields $\sum_{i=1}^{k} \alpha_i v_i - \sum_{i=1}^{n-k} \beta_i u_i = 0$. Since $\{v_1, \ldots, v_k, u_1, \ldots, u_{n-k}\}$ is a basis of $\mathcal{V}$ and, therefore, linearly independent, it follows that $\alpha_1 = \ldots = \alpha_k = \beta_1 = \ldots = \beta_{n-k} = 0$ and, consequently, that $v = 0$. □

LEMMA 3.5. *With T and $\mathcal{M}$ as above, $T : \mathcal{M} \to \mathcal{R}(T)$ is 1-1 and onto.*

Proof. If $v \in \mathcal{M}$ is such that $T(v) = 0$, then $v \in \mathcal{N}(T) \cap \mathcal{M}$, which is $\{0\}$ by Proposition 3.4. It follows from Proposition 3.2 that T is 1-1 on $\mathcal{M}$. If $w \in \mathcal{R}(T)$, then $w = T(v)$ for some $v \in \mathcal{V}$. With the basis $\{v_1, \ldots, v_k, u_1, \ldots, u_{n-k}\}$ as above, we can write $v = \sum_{i=1}^{k} \alpha_i v_i + \sum_{i=1}^{n-k} \beta_i u_i$ for some scalars α_i and β_i. Since $\sum_{i=1}^{k} \alpha_i v_i \in \mathcal{N}(T)$, we have that $w = T(v) = T\left(\sum_{i=1}^{k} \alpha_i v_i\right) + T\left(\sum_{i=1}^{n-k} \beta_i u_i\right) = T\left(\sum_{i=1}^{n-k} \beta_i u_i\right)$. Since $\sum_{i=1}^{n-k} \beta_i u_i \in \mathcal{M}$, it follows that T maps $\mathcal{M}$ onto $\mathcal{R}(T)$. □

THEOREM 3.6. *If $\mathcal{V}$ is finite-dimensional with $\dim \mathcal{V} = n$ and $T : \mathcal{V} \to \mathcal{W}$ is a linear transformation, then $\dim \mathcal{N}(T) + \dim \mathcal{R}(T) = n$.*

Proof. Suppose that, as above, we have a basis $\{v_1, \ldots, v_k\}$ for $\mathcal{N}(T)$, an expanded basis $\{v_1, \ldots, v_k, u_1, \ldots, u_{n-k}\}$ for $\mathcal{V}$, and $\mathcal{M} = \operatorname{span}\{u_1, \ldots, u_{n-k}\}$. Since $\{u_1, \ldots, u_{n-k}\}$ is a basis for $\mathcal{M}$, we have that $\dim \mathcal{N}(T) + \dim \mathcal{M} = n$. We also have, by Proposition 3.2 and Lemma 3.5, that $\{T(u_1), \ldots, T(u_{n-k})\}$ is a linearly independent spanning set in $\mathcal{R}(T)$, i.e., a basis for $\mathcal{R}(T)$. It follows that $\dim \mathcal{R}(T) = \dim \mathcal{M} = n - k$ and, hence, that $\dim \mathcal{N}(T) + \dim \mathcal{R}(T) = n$. □

Linear equations.

A *linear equation* has the form $T(v) = w$, where $T : \mathcal{V} \to \mathcal{W}$ is a given linear transformation, $w \in \mathcal{W}$ is a given right-hand side, and $v \in \mathcal{V}$ is a solution to be determined. Our immediate goal is to look into the existence and uniqueness of solutions.

Clearly, $T(v) = w$ has a solution if and only if $w \in \mathcal{R}(T)$. The following somewhat refined statement is also clear.

PROPOSITION 3.7.

(a) *$T(v) = w$ has at least one solution for every $w \in \mathcal{W}$ if and only if T is onto, i.e., $\mathcal{R}(T) = \mathcal{W}$.*

(b) *$T(v) = w$ has at most one solution for every $w \in \mathcal{W}$ if and only if T is one-to-one, i.e., $\mathcal{N}(T) = \{0\}$.*

Suppose that $\mathcal{V}$ and $\mathcal{W}$ are finite-dimensional with $\dim \mathcal{V} = n$ and $\dim \mathcal{W} = m$. We consider the cases $n > m$, $n < m$, and $n = m$ in order, recalling from Theorem 3.6 that $\dim \mathcal{N}(T) + \dim \mathcal{R}(T) = n$.

If $n > m$, then $\dim \mathcal{N}(T) = n - \dim \mathcal{R}(T) \geq n - m > 0$. It follows that if $w \in \mathcal{R}(T)$, then there are *infinitely many solutions* of $T(v) = w$. Indeed, we have the following characterization of solutions, which is also valid when $n \leq m$ and even when $\mathcal{V}$ or $\mathcal{W}$ is infinite-dimensional.

Proposition 3.8. *If $v \in \mathcal{V}$ is any vector satisfying $T(v) = w$, then the set of all solutions is* $\{v + u : \ u \in \mathcal{N}(T), \text{ i.e., } T(u) = 0\}$.

Proof. If $u \in \mathcal{N}(T)$, then $T(v + u) = T(v) + T(u) = T(v) + 0 = w$; thus everything in the set is a solution. If $s \in \mathcal{V}$ is any vector satisfying $T(s) = w$, then $T(s - v) = T(s) - T(v) = w - w = 0$. Consequently, $u = s - v \in \mathcal{N}(T)$, and $s = v + u$ is in the set. $\square$

Since $\dim \mathcal{N}(T) > 0$ when $n > m$, it follows immediately that if a solution exists, then there are infinitely many other solutions. Note that a solution exists for every $w \in \mathcal{W}$ if and only if $\dim \mathcal{N}(T) = n - m$.

If $n < m$, then $\dim \mathcal{R}(T) = n - \dim \mathcal{N}(T) \leq n < m$. It follows that T cannot be onto and that there are *no solutions* of $T(v) = w$ for some $w \in \mathcal{W}$.

In summary, if $n > m$, then a solution of $T(v) = w$ *may exist* for every $w \in \mathcal{W}$, but it is *never unique*. If $n < m$, then a solution *may be unique* if it exists, but there are some $w \in \mathcal{W}(T)$ for which a solution *does not exist*.

So our only hope for both existence and uniqueness of a solution of $T(v) = w$ for every $w \in \mathcal{W}$ lies in the final case $m = n$.

Theorem 3.9. *Suppose that $T : \mathcal{V} \to \mathcal{W}$ is linear and that $\mathcal{V}$ and $\mathcal{W}$ are finite-dimensional with $\dim \mathcal{V} = \dim \mathcal{W}$. Then the following are equivalent:*

(a) *$T(v) = w$ has a unique solution $v \in \mathcal{V}$ for every $w \in \mathcal{W}$.*

(b) *T is 1-1, i.e., $T(v) = w$ has at most one solution for every $w \in \mathcal{W}$.*

(c) *T is onto, i.e., $T(v) = w$ has at least one solution for every $w \in \mathcal{W}$.*

(d) *T has an inverse, i.e., a linear transformation $T^{-1} : \mathcal{W} \to \mathcal{V}$ such that $T^{-1}(T(v)) = v$ for every $v \in \mathcal{V}$ and $T(T^{-1}(w)) = w$ for every $w \in \mathcal{W}$.*

Proof.
(a)$\Rightarrow$(b) If (a) holds, then $T(v) = 0$ has the unique solution $v = 0$, and it follows from Proposition 3.2 that T is 1-1.

(b)$\Rightarrow$(c) If T is 1-1, then $\dim \mathcal{N}(T) = 0$. Since $\dim \mathcal{N}(T) + \dim \mathcal{R}(T) = n$ by Theorem 3.6, it follows that $\dim \mathcal{R}(T) = n$ and, consequently, that T is onto.

(c)$\Rightarrow$(a) If T is onto, then $T(v) = w$ has at least one solution $V \in \mathcal{V}$ for every $w \in \mathcal{W}$. Also, we have that $\mathcal{R}(T) = \mathcal{W}$, and so $\dim \mathcal{R}(T) = n$. Since $\dim \mathcal{N}(T) + \dim \mathcal{R}(T) = n$ by Theorem 3.6, it follows that $\dim \mathcal{N}(T) = 0$. Then, by Proposition 3.1, T is 1-1, and the solution of $T(v) = w$ is unique for every $w \in \mathcal{W}$.

To complete the proof, recall from Proposition 3.1 that T is 1-1 and onto if and only if it has an inverse. Thus (a)-(c) hold if and only if (d) holds. $\square$

Linear equations in $\mathbb{R}^{\mathrm{n}}$.

Here, we outline the implications of the general results above when $\mathcal{V} = \mathbb{R}^{\mathrm{n}}$, $\mathcal{W} = \mathbb{R}^{\mathrm{m}}$, and $T(v) = Av \in \mathbb{R}^{\mathrm{m}}$ for $v \in \mathbb{R}^{\mathrm{n}}$, where $A \in \mathbb{R}^{\mathrm{m}\times\mathrm{n}}$. One can regard T as defined by $T(v) = Av$ for a given A, or one can regard A as the matrix representation of a given T with respect to the natural bases on $\mathbb{R}^{\mathrm{n}}$ and $\mathbb{R}^{\mathrm{m}}$. We denote $\mathcal{R}(T)$ by $\mathcal{R}(A)$ and $\mathcal{N}(T)$ by $\mathcal{N}(A)$.

In this context, we write a linear equation as $Ax = b$, where $b \in \mathbb{R}^{\mathrm{m}}$ is given and $x \in \mathbb{R}^{\mathrm{n}}$ is a solution to be determined. This represents a system of m linear equations in n unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned}$$

From our general results, we have the following:

1. $\dim \mathcal{N}(A) + \dim \mathcal{R}(A) = n$.
2. If $Ax = b$ for some $x \in \mathbb{R}^{\mathrm{n}}$, then the set of *all* solutions is $\{x + y : \ y \in \mathcal{N}(A), \text{ i.e., } Ay = 0\}$.
3. If $n > m$, then a solution *may exist* for every $b \in \mathbb{R}^{\mathrm{m}}$, but it is *never unique*. If $n < m$, then a solution *may be unique* if it exists, but there are some $b \in \mathbb{R}^{\mathrm{m}}$ for which a solution *does not exist*.
4. If $n = m$, then the following are equivalent:
   (a) $Ax = b$ has a unique solution for every $b \in \mathbb{R}^{\mathrm{n}}$.
   (b) The only solution of $Ax = 0$ is $x = 0$.
   (c) $Ax = b$ has at least one solution for every $b \in \mathbb{R}^{\mathrm{n}}$.
   (d) A has an *inverse matrix* $A^{-1} \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}}$ such that $A^{-1}A = AA^{-1} = I$, the identity matrix[1] in $\mathbb{R}^{\mathrm{n}\times\mathrm{n}}$.

Items 1-3 are straightforward “translations” of Theorem 3.6, Proposition 3.8, and the summary preceding Theorem 3.9, respectively. Parts (a) and (c) of the item 4 are likewise straightforward “translations” of their counterparts in Theorem 3.9. Part (b) is nearly straightforward but requires recalling from Proposition 3.2 that the linear transformation defined by A is 1-1 if and only if $\mathcal{N}(A) = \{0\}$, which is to say that the only solution of $Ax = 0$ is $x = 0$. For part (d), recall from Proposition 3.1 that part (a) holds if and only if the map T associated with A has an inverse T^{-1} such that $T^{-1}(T(x)) = T(T^{-1}(x)) = x$ for every $x \in \mathbb{R}^{\mathrm{n}}$. Then take $A^{-1} \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}}$ to be the matrix representation of T^{-1} with respect to the natural basis on $\mathbb{R}^{\mathrm{n}}$. It is easy to verify that $A^{-1}Ax = AA^{-1}x = x$ for all $x \in \mathbb{R}^{\mathrm{n}}$, which is to say that $A^{-1}A = AA^{-1} = I$.

[1]This is the matrix with all diagonal entries equal to one and all other entries equal to zero.

Remark 1. If $A \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}}$ has an inverse matrix A^{-1}, then A is said to be *invertible* or *nonsingular*. Note that $Ax = b \iff A^{-1}Ax = A^{-1}b \iff x = A^{-1}b$. This characterization of the solution is useful for theoretical purposes but does not suggest a practical way to determine x in general.

PROPOSITION 3.10. *Suppose A, $B \in \mathbb{R}^{\mathrm{n}\times\mathrm{n}}$. Then AB is nonsingular if and only if both A and B are nonsingular.*

Remark 2. In the present context, we can specify an additional equivalent condition in item 4, as follows: Denote the columns of A by

$$a_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}, \quad a_2 = \begin{pmatrix} a_{12} \\ \vdots \\ a_{n2} \end{pmatrix}, \quad \ldots, \quad a_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix}.$$

Then for $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{\mathrm{n}}$, we have $Ax = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{nj}x_j \end{pmatrix} = x_1a_1 + x_2a_2 + \ldots + a_nx_n$.

Letting x range over all $\mathbb{R}^{\mathrm{n}}$, we have that $\mathcal{R}(A) = \operatorname{span}\{a_1, \ldots, a_n\}$. It follows that A is onto $\iff$ $\mathcal{R}(A) = \mathbb{R}^{\mathrm{n}} \iff \operatorname{span}\{a_1, \ldots, a_n\} = \mathbb{R}^{\mathrm{n}} \iff$ the columns of A are linearly independent.

DEFINITION 3.11. *The column rank of a matrix is the maximal number of linearly independent columns in the matrix.*

We have that the column rank of A is equal $\dim \operatorname{span}\{a_1, \ldots, a_n\}$. Then to item 4 above, we can add a fifth equivalent condition:

(e) The column rank of A is n.