# TLDR

Quick summary of today's notes. Lecture starts on next page.

- Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear transformation with standard matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

  If $v \in \mathbb{R}^2$, then $T(v) = Av$ is the vector in $\mathbb{R}^2$ formed by rotating $v$ counterclockwise by $\theta$ radians.

- A function $f : X \to Y$ is *one-to-one* (or *injective*) if we never have $f(a) = f(b)$ for $a \neq b$.
- A function $f : X \to Y$ is *onto* (or *surjective*) if for each $y \in Y$, there exists $x \in X$ with $f(x) = y$.
- Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation with standard matrix $A$.

  Then $T$ is one-to-one if and only if the columns of $A$ are linearly independent.

  This happens if and only if every column of $A$ contains a pivot position.

  Likewise, $T$ is onto if and only if the span of the columns of $A$ is $\mathbb{R}^m$.

  This happens if and only if every row of $A$ contains a pivot position.

- If $T$ and $U$ are functions $\mathbb{R}^n \to \mathbb{R}^m$, then there is a natural way to form the sum $T + U$ and the scalar multiple $cT$ for $c \in \mathbb{R}$. The definitions are

$$(T + U)(v) = T(v) + U(v) \qquad \text{and} \qquad (cT)(v) = c \cdot T(v) \qquad \text{for } v \in \mathbb{R}^n.$$

  Both of these are also functions $\mathbb{R}^n \to \mathbb{R}^m$.

  If $T$ and $U$ are linear, then $T + U$ and $cT$ are both linear.

- If $A$ and $B$ are $m \times n$ matrices, then there is a natural way to form the sum $A + B$ and the scalar multiple $cA$ for $c \in \mathbb{R}$. Both of these are also $m \times n$ matrices.
- If $A$ and $B$ are the standard matrices of $T$ and $U$, then $A + B$ is the standard matrix of $T + U$, and $cA$ for $c \in \mathbb{R}$ is the standard matrix of $cT$.

# 1 Last time: linear transformations

A *function* is a rule transforms that inputs from one set to outputs in another set.

Saying that $f : X \to Y$ is a function means that $f$ is a rule that transforms inputs in $X$ to outputs in $Y$.

The set $X$ is called the *domain* of $f$ and the set $Y$ is called the *codomain* of $f$.

Let $m$ and $n$ be positive integers. Recall that $\mathbb{R}^n$ is the set of vectors with $n$ rows.

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a function, whose domain and codomain are the sets of vectors $\mathbb{R}^n$ and $\mathbb{R}^m$.

The following mean the same thing:

- $T$ is *linear* is the sense that $T(u + v) = T(u) + T(v)$ and $T(cv) = cT(v)$ for $u, v \in \mathbb{R}^n$, $c \in \mathbb{R}$.
- There is an $m \times n$ matrix $A$ such that $T$ has the formula $T(v) = Av$ for $v \in \mathbb{R}^n$.

If we are given a linear transformation $T$, then $T(v) = Av$ for the matrix

$$A = \begin{bmatrix} T(e_1) & T(e_2) & \dots & T(e_n) \end{bmatrix}$$

where $e_i \in \mathbb{R}^n$ is the vector with a 1 in row $i$ and 0 in all other rows.

We call $A$ the *standard matrix* of $T$.

Two different linear functions $\mathbb{R}^n \to \mathbb{R}^m$ cannot have the same standard matrix.

**Example.** Fix $\theta \in [0, 2\pi)$. The notation $[a, b)$ means "the set of numbers $x \in \mathbb{R}$ with $a \le x < b$." Define

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

and let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation $T(v) = Av$.

If $v = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is a vector parallel to the $x$-axis, then $T(v) = Av = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$.

If $v = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is a vector parallel to the $y$-axis, then $T(v) = Av = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = \begin{bmatrix} \cos(\theta + \frac{\pi}{2}) \\ \sin(\theta + \frac{\pi}{2}) \end{bmatrix}$.

In general, $T(v) = Av$ is the vector obtained by rotating $v$ counterclockwise by the angle $\theta$.

This holds since any vector $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ can be written $v = \begin{bmatrix} v_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ v_2 \end{bmatrix}$, so is the arrow to the opposite vertex in the parallelogram with sides $\begin{bmatrix} v_1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ v_2 \end{bmatrix}$.

Since $T(v) = T\left(\begin{bmatrix} v_1 \\ 0 \end{bmatrix}\right) + T\left(\begin{bmatrix} 0 \\ v_2 \end{bmatrix}\right)$ and since $T$ rotates by angle $\theta$ the two vectors on the right, $T(v)$ is the arrow from 0 to the opposite vertex in our previous parallelogram, now rotated by angle $\theta$.

## 2 Geometric interpretations of linear transformations $\mathbb{R}^2 \to \mathbb{R}^2$

Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation with standard matrix $A$. We can illustrate $T$ by drawing the parallelogram with sides $T(e_1)$ and $T(e_2)$. (Fill in these pictures yourself.)

| Standard matrix of $T$ | Picture | Description of $T$ |
|---|---|---|
| $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ | | Reflect across the $x$-axis |
| $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ | | Reflect across the $y$-axis |
| $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ | | Reflect across $y = x$ |
| $\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$ $(0 < k < 1)$ | | Horizontal contraction |
| $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$ $(0 < k < 1)$ | | Vertical contraction |
| $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ $(k > 0)$ | | Horizontal sheering |

# 3 One-to-one and onto functions

This section talks about two important classes of linear transformations, which can be characterized in terms of whether the columns of the standard matrix are linearly independent or span the codomain.

**Definition.** A function $f : X \to Y$ is *one-to-one* (or *injective*) if $f(a) = f(b)$ implies $a = b$.

This means that $f$ does not send two different inputs to the same output.

If $a \neq b$ and $f(a) = f(b)$ then $f$ is not one-to-one.

**Example.** Suppose $T : \mathbb{R}^3 \to \mathbb{R}^2$ is the linear transformation $T(v) = Av$ where

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 5 & 3 \end{bmatrix}.$$

Is $T$ one-to-one? No: since $A$ has more columns than rows, its columns are linearly dependent. Therefore there is a vector $0 \neq v \in \mathbb{R}^3$ such that $T(v) = Av = 0$. But we also have $T(0) = 0$.

**Theorem.** If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation then the following mean the same thing:

(a) $T$ is one-to-one.

(b) The only solution to $T(x) = 0$ is $x = 0 \in \mathbb{R}^n$.

(c) The columns of the standard matrix $A$ of $T$ are linearly independent.

*Proof.* Suppose the only solution to $T(x) = 0$ is $x = 0 \in \mathbb{R}^n$. Then whenever $u, v \in \mathbb{R}^n$ are vectors with $u \neq v$, we have $T(u) - T(v) = T(u - v) \neq 0$ since $u - v \neq 0$, so $T(u) \neq T(v)$. Therefore $T$ is one-to-one.

If $T$ is one-to-one, then $T(x) = T(0) = 0$ implies $x = 0$, so the only solution to $T(x) = 0$ is $x = 0$. □

**Definition.** A function $f : X \to Y$ is *onto* (or *surjective*) if $\text{range}(f) = \{f(x) : x \in X\} = Y$.

This mean that the range of $f$ is equal to its codomain.

If there is a value $y \in Y$ such that $f(x) \neq y$ for all $x \in X$, then $f$ is not onto.

**Example.** Suppose again that $T : \mathbb{R}^3 \to \mathbb{R}^2$ is the linear transformation $T(v) = Av$ where

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 5 & 3 \end{bmatrix}.$$

Is $T$ onto? Yes: the span of the columns of $A$ is $\mathbb{R}^2$ if and only if $A$ has a pivot in every row, and

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 5 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 3/5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 19/5 \\ 0 & 1 & 3/5 \end{bmatrix} = \mathsf{RREF}(A).$$

**Theorem.** Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is the linear transformation with standard matrix $A$.

The following properties are all equivalent:

(a) $T$ is onto.

(b) The matrix equation $Ax = b$ has a solution for each $b \in \mathbb{R}^m$

(c) The span of the columns of $A$ of $T$ is $\mathbb{R}^m$.

*Proof.* The vectors in the range of $T$ are precisely the linear combinations of the columns of $A$.

This is $\mathbb{R}^m$ precisely when the span of the columns of $A$ is $\mathbb{R}^m$. □

**Example.** Suppose $T : \mathbb{R}^2 \to \mathbb{R}^3$ is the function $T\left(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right) = \begin{bmatrix} 3v_1 + v_2 \\ 5v_1 + 7v_2 \\ v_1 + 3v_2 \end{bmatrix}$.

This function is a linear transformation. Its standard matrix is $A = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix}$.

To determine if $T$ is one-to-one, we check if the columns of $A$ are linearly independent. To do this, we convert $A$ to its reduced echelon form:

$$A = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 \\ 5 & 7 \\ 3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 \\ 0 & -8 \\ 0 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \mathsf{RREF}(A).$$

This shows that $A$ has a pivot position in every column, which means that the only solution to $Ax = 0$ is $x = 0$, which means the columns of $A$ are linearly independent, which means $T$ is one-to-one.

To determine if $T$ is onto, we want to find out if the columns of $A$ span $\mathbb{R}^3$. From last time, we know that this happens if and only if $A$ has a pivot position in every row. Since the third row of $A$ has no pivot position, $T$ is not onto.

**Corollary.** A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one only if $n \leq m$, and onto only if $n \geq m$.

*Proof.* Results last time show that $T$ is one-to-one if and only if its standard matrix has a pivot position in every column, and onto if and only if its standard matrix has a pivot position in every row. The first case requires there to be more columns $n$ than rows $m$, and the second case requires there to be more rows $m$ than columns $n$ (since each row and each column contains at most one pivot position). □

There is a more efficient way to state the conditions that characterize when a linear transformation is one-to-one or onto:

**Theorem.** Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation with standard matrix $A$.

(a) The linear transformation $T$ is one-to-one if and only if the matrix $A$ has a pivot in every column.

(b) The linear transformation $T$ is onto if and only if the matrix $A$ has a pivot in every row.

*Proof.* We have seen that $T$ is one-to-one if and only if the columns of $A$ are linearly independent, which occurs precisely when $A$ has a pivot in every column.

Likewise, we have seen that $T$ is onto if and only if the span of the columns of $A$ is $\mathbb{R}^m$, which occurs precisely when $A$ has a pivot in every row. □

# 4 Operators on linear transformations and matrices

There are several simple, natural operations we can use to combine and alter linear transformations to get other linear transformations. Our next goal is to translate these operations into matrix operations.

**Sums and scalar multiples**. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ and $U : \mathbb{R}^n \to \mathbb{R}^m$ are two linear functions with the same domain and codomain. Their sum $T + U$ is the function $\mathbb{R}^n \to \mathbb{R}^m$ defined by

$$(T + U)(v) = T(v) + U(v) \qquad \text{for } v \in \mathbb{R}^n.$$

If $c \in \mathbb{R}$ is a scalar, then $cT$ is the function $\mathbb{R}^n \to \mathbb{R}^m$ defined by

$$(cT)(v) = cT(v) \qquad \text{for } v \in \mathbb{R}^n.$$

**Fact.** Both $T+U$ and $cT$ are linear transformations.

*Proof.* To see that $T+U$ is linear, we check that

$$(T+U)(u+v) = T(u+v) + U(u+v) = T(u) + T(v) + U(u) + U(v) = (T+U)(u) + (T+U)(v)$$

for $u, v \in \mathbb{R}^n$, and $(T+U)(av) = T(av) + U(av) = aT(v) + aU(v) = a(T+U)(v)$ for $a \in \mathbb{R}$ and $v \in \mathbb{R}^n$. Since these properties hold, $T+U$ is linear. The proof that $cT$ is linear is similar. (Try this yourself!) □

Since sums and scalar multiples of linear functions are linear, it follows that differences $T-U$ and arbitrary linear combinations $aT + bU + cV + \dots$ of linear functions are linear.

Suppose $T$ and $U$ have standard matrices

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & & b_{2n} \\ \vdots & & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

so that $T(v) = Av$ and $U(v) = Bv$.

**Proposition.** The standard matrices of $T+U$ and $cT$ are

$$A+B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & & a_{2n}+b_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix} \quad \text{and} \quad cA = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1n} \\ ca_{21} & ca_{22} & & ca_{2n} \\ \vdots & & \ddots & \vdots \\ ca_{m1} & ca_{m2} & \dots & ca_{mn} \end{bmatrix}.$$

This is how we *define* sums and scalar multiples of matrices. These operations work in essentially the same way as for vectors: we can add matrices of the same size, by adding the entries in corresponding positions together, and we can multiply a matrix by a scalar $c$ by multiplying all entries by $c$.

**Example.** We have

$$\begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}.$$

and

$$-\begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} + 2\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} -4 & 0 & -5 \\ 1 & -3 & -2 \end{bmatrix} + \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} -2 & 2 & -3 \\ 7 & 7 & 12 \end{bmatrix}.$$

Suppose $T, U, V$ are linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ with standard matrices $A, B, C$. Let $a, b \in \mathbb{R}$.

The following properties then hold:

| Functions | Matrices |
| --- | --- |
| 1. $T + U = U + V$ | $A + B = B + A$. |
| 2. $(T+U)+V = T+(U+V)$ | $(A+B)+C = A+(B+C)$. |
| 3. $T + 0 = T$ where $0 : \mathbb{R}^n \to \mathbb{R}^m$ is the map $0(v) = 0 \in \mathbb{R}^m$. | $A + 0 = A$. |
| 4. $a(T+U) = aT + aU$ | $a(A+B) = aA + aB$. |
| 5. $(a+b)T = aT + bT$ | $(a+b)A = aA + bA$. |
| 6. $a(bT) = (ab)T$. | $a(bA) = (ab)A$. |

# 5 Vocabulary

Keywords from today's lecture:

1. **One-to-one** or **injective** function $f : X \to Y$.

   A function with the property that if $f(u) = f(v)$ for $u, v \in X$ then $u = v$.

   Example: The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$.

   The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is not one-to-one: $f(-2) = f(2) = 4$.

2. **Onto** or **surjective** function $f : X \to Y$.

   A function with the property that $y \in Y$ then there exists $x \in X$ with $f(x) = y$.

   Example: The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$.

   The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is not onto: no negative number is in its range.

3. **Sums** and **scalar multiples** of linear functions.

   If $T : \mathbb{R}^n \to \mathbb{R}^m$ and $U : \mathbb{R}^n \to \mathbb{R}^m$ and $c \in \mathbb{R}$ then

   $$T + U : \mathbb{R}^n \to \mathbb{R}^m$$

   is the function with $(T + U)(v) = T(v) + U(v)$, and

   $$cT : \mathbb{R}^n \to \mathbb{R}^m$$

   is the function with $(cT)(v) = c(T(v))$.

4. **Sums** and **scalar multiples** of matrices.

   If $A$ and $B$ are $m \times n$ matrices then $A + B$ is the $m \times n$ matrix whose entry in position $(i, j)$ is $A_{ij} + B_{ij}$. If $c \in \mathbb{R}$ then $cA$ is the matrix whose entry in position $(i, j)$ is $cA_{ij}$.