This document is an **exact transcript** of the lecture, with extra summary and vocabulary sections. Due to time constraints, the lectures sometimes only contain limited illustrations, proofs, and examples. For a more thorough discussion of the course content, **consult the textbook**.

Summary

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

$$A = \left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right].$$

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

1 Last time: linear transformations

Writing $f : X \to Y$ means that X is the domain (the set of inputs), Y is the codomain (a set that contains all outputs and possibly more elements), and f is a function that transforms each input in X to an output that belongs to Y.

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

$$A = \left[\begin{array}{cccc} T(e_1) & T(e_2) & \dots & T(e_n) \end{array}\right]$$

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 $a \leq x < b$." Define

$$A = \left[\begin{array}{rr} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right]$$

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

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

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

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

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

As $T(v) = T\left(\left[\begin{array}{c} v_1 \\ 0 \end{array}\right]\right) + T\left(\left[\begin{array}{c} 0 \\ v_2 \end{array}\right]\right)$ and as T rotates by angle θ the vectors $\left[\begin{array}{c} v_1 \\ 0 \end{array}\right]$ and $\left[\begin{array}{c} 0 \\ v_2 \end{array}\right]$, the vector $T(v)$ is the arrow from 0 to the opposite vertex in our parallelogram, now rotated by angle θ.

2 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 standard matrix has pivots in every row or in every column. To prepare for this, we first discuss some properties of functions in general.

The definition of a function involves **three components**: a choice of **domain**, a choice of **codomain**, and a **formula/rule/algorithm** to assign an output in the codomain to each input in the domain. Two functions are the same only when all three of these components are equal.

For example, the formula $f(x) = \sqrt{x}$ defines a function $X \to Y$ with $X = \{x \in \mathbb{R} : x \geq 0\}$ and $Y = \mathbb{R}$.

The formula $f(x) = \sqrt{x}$ also defines a function $X \to X$ with $X = \{x \in \mathbb{R} : x \geq 0\}$. We consider this to be a different function from the previous example, because it has a different codomain.

Likewise, the formula $f(x) = |x|$ defines a function $\mathbb{R} \to \mathbb{R}$ and a different function $\mathbb{R} \to \{x \in \mathbb{R} : x \geq 0\}$.

For every x in the domain X of f, we get an output $f(x)$.

It is possible that some values y in the codomain Y may never occur as outputs of f.

The *image* of an input x in X under f is the output $f(x)$.

The *range* of the function f (also called the *image* of f) is the set $\mathsf{range}(f) = \{f(x) : x \in X\}$ of images of all inputs in the domain. This is the subset of the codomain Y that gives all actual outputs of f.

Example. Suppose $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$ and $T : \mathbb{R}^2 \to \mathbb{R}^3$ is the function defined by $T(v) = Av$.

(a) The image of a single vector $v \in \mathbb{R}^2$ under T is by definition $T(v) = Av$.

The image of $v = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ under T is $T\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$.

(b) Is the range of T all of $\mathbb{R}^3$?

No: if it was, then the span of the columns of A would be all of $\mathbb{R}^3$, which happens precisely when A has a pivot position in every row. This is impossible since each column can only contain one pivot position, but A has three rows and only two columns. Therefore $\mathsf{range}(T) \neq \mathbb{R}^3$.

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.

The function f is **not** one-to-one if there are different inputs $a \neq b$ with $f(a) = f(b)$.

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

The following properties are all equivalent:

(a) T is one-to-one.

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

(c) The columns of A are linearly independent.

(d) The matrix A of T has a pivot position in every column.

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 $x = 0$.

This shows that (a) and (b) are equivalent. Properties (b) and (c) are equivalent by definition, and we saw in an earlier lecture that (c) and (d) are equivalent. □

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

Thus, f is onto if its range is equal to its codomain.

The function f is **not** onto if there is a value $y \in Y$ such that $f(x) \neq y$ for all $x \in X$.

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 a 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 is $\mathbb{R}^m$.

(d) The matrix A has a pivot position in every row.

Proof. Property (a) means that $T(x) = b$ has a solution for each $b \in \mathbb{R}^m$, which is the same thing as (b).

We saw in an earlier lecture that (b), (c), and (d) are equivalent. □

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 linear with standard matrix $A = \begin{bmatrix} T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) & T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix}$.

To determine if T is one-to-one, we check if every column of A contains a pivot position.

To do this, we convert A to its reduced echelon form:

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

Evidently A does have a pivot position in every column, so T is one-to-one.

To determine if T is onto, we want to find out if A contains a pivot position in every row.

Since the third row of $\mathsf{RREF}(A)$ is zero, 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. Let A be the standard matrix of T. Then T is one-to-one if and only if A has a pivot position in every column, and T is onto if and only if A has a pivot position in every row.

Each row and each column contains at most one pivot position. Thus if A has a pivot in every column then the number of columns n cannot be more than the number of rows m. Likewise, if A has a pivot in every row then the number of rows m cannot be more than the number of columns n.

This means that if T is one-to-one then $n \leq m$ and if T is onto then $m \leq n$. $\square$

3 Vocabulary

Keywords from today's lecture:

1. **Range** of a function $f : X \to Y$.

   The set $\mathsf{range}(f) = \{f(x) : x \in X\} \subset Y$ of all actual outputs of the function f.

   Example: If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $T : \mathbb{R}^3 \to \mathbb{R}^3$ has $T(v) = Av$ then $\mathsf{range}(T) = \left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} : x, y \in \mathbb{R} \right\}$.

2. **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: $f(-2) = f(2) = 4$.

3. **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.