

This document is an **exact transcript** of the lecture, with extra summary and vocabulary sections for your convenience. 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.

Matrix-vector products:

- We can multiply an $m \times n$ matrix A by a vector $v \in \mathbb{R}^n$. The result, written Av , belongs to \mathbb{R}^m .

If $a_1, a_2, \dots, a_n \in \mathbb{R}^m$ are the columns of A and $v_1, v_2, \dots, v_n \in \mathbb{R}$ are the entries of v then

$$Av = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1 a_1 + v_2 a_2 + \dots + v_n a_n.$$

Here is a concrete example:

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ -1 & 0 & -2 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 10 \\ 1000 \\ 10000 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 10 \begin{bmatrix} 2 \\ 0 \end{bmatrix} + 1000 \begin{bmatrix} 3 \\ -2 \end{bmatrix} + 10000 \begin{bmatrix} 4 \\ -3 \end{bmatrix} = \begin{bmatrix} 4321 \\ -3201 \end{bmatrix}.$$

- If A is $m \times n$, $u, v \in \mathbb{R}^n$, and $c \in \mathbb{R}$ then $A(u + v) = Au + Av \in \mathbb{R}^m$ and $A(cv) = c(Av) \in \mathbb{R}^m$. We say that $v \mapsto Av$ (the function whose output, given input $v \in \mathbb{R}^n$, is $Av \in \mathbb{R}^m$) is **linear**.

Matrix equations:

- If A is an $m \times n$ matrix, $b \in \mathbb{R}^m$, and $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a vector of n variables, then $Ax = b$ is a **matrix equation**. This has the same solutions as the linear system with augmented matrix $\begin{bmatrix} A & b \end{bmatrix}$.
- The matrix equation $Ax = b$ has a solution for all b if and only if A has a pivot in every row.

Linear independence:

- Vectors $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ are **linearly independent** if the only way to express $0 \in \mathbb{R}^n$ as a linear combination $c_1 v_1 + c_2 v_2 + \dots + c_p v_p$ for $c_1, c_2, \dots, c_p \in \mathbb{R}$ is by taking $c_1 = c_2 = \dots = c_p = 0$.

Vectors that are not linearly independent are **linearly dependent**.

- Any sufficiently large set of vectors in \mathbb{R}^n is linearly dependent.

Specifically, if $p > n$ then any vectors $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ are linearly dependent.

1 Last time: Vectors

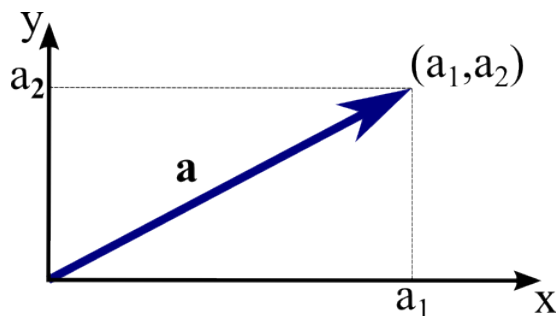
A (*column*) *vector* of size n is an $n \times 1$ matrix: $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$.

Let \mathbb{R}^n be the set of all vectors with exactly n rows.

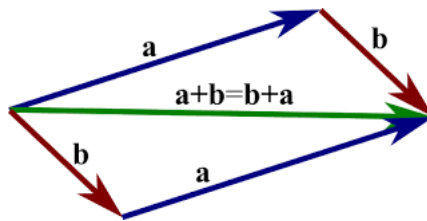
We can add two vectors of the same size: $\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$.

We can multiply a vector by a *scalar*: $cv = c \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}$ for $c \in \mathbb{R}$ and $v \in \mathbb{R}^n$.

We visualize vectors $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \in \mathbb{R}^2$ as arrows in the Cartesian plane from the origin to $(x, y) = (a_1, a_2)$:



Relative to this picture, the sum $a + b$ of two vectors $a, b \in \mathbb{R}^2$ is the vector represented by the arrow from the origin to the point which is the opposite vertex of the parallelogram with sides a and b :



The *zero vector* $0 \in \mathbb{R}^n$ is the vector $0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$. We have $0 + v = v + 0 = v$ for any vector v .

A *linear combination* of vectors $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ is any vector of the form

$$y = c_1 v_1 + c_2 v_2 + \dots + c_p v_p \in \mathbb{R}^n \text{ for any choice of numbers } c_1, c_2, \dots, c_p \in \mathbb{R}.$$

The *span* of some vectors $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ is the set of all of their linear combinations. Denote this by

$$\mathbb{R}\text{-span}\{v_1, v_2, \dots, v_p\} \quad \text{or} \quad \text{span}\{v_1, v_2, \dots, v_p\}.$$

In terms of geometry, the span of a set of vectors in \mathbb{R}^2 is either a point (at the origin), a line (through the origin), or the whole plane \mathbb{R}^2 . The span of a set of vectors in \mathbb{R}^3 is either a point (at the origin), a line (through the origin), a plane (containing the origin), or all of \mathbb{R}^3 .

Proposition. If $v_1, v_2, \dots, v_p \in \mathbb{R}^n$, then a vector $y \in \mathbb{R}^n$ belongs to $\mathbb{R}\text{-span}\{v_1, v_2, \dots, v_p\}$ if and only if the $n \times (p+1)$ matrix $\begin{bmatrix} v_1 & v_2 & \dots & v_p & y \end{bmatrix}$ is the augmented matrix of a consistent linear system.

The notation $\begin{bmatrix} v_1 & v_2 & \dots & v_p & y \end{bmatrix}$ means the matrix whose i th column is v_i and last column is y .

2 Multiplying matrices and vectors

So far we have been using matrices as a compact notation for representing linear systems.

Today we introduce a second, perhaps more fundamental way of viewing a matrix: namely, as an operator that transforms one vector to another.

Definition. If A is a matrix with columns $a_1, a_2, \dots, a_n \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$, so that

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

then the *matrix-vector product* Av is the vector in \mathbb{R}^m given by:

$$Av = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = v_1 a_1 + v_2 a_2 + \dots + v_n a_n \in \mathbb{R}^m.$$

Thus Av is the linear combination of the columns of A with coefficients given by the entries of v .

Example. If $A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}$ and $v = \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix}$ then $a_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $a_2 = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$, and $a_3 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ so

$$Av = 4a_1 + 3a_2 + 7a_3 = \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

Example. If $A = \begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix}$ and $v = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$ then $a_1 = \begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix}$ and $a_2 = \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}$ so we have

$$Av = 4a_1 + 7a_2 = 4 \begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7 \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix} + \begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}.$$

If A is $m \times n$ then Av is only defined for $v \in \mathbb{R}^n$ (when v is an $n \times 1$ matrix), and in this case $Av \in \mathbb{R}^m$.

Thus A transforms vectors in \mathbb{R}^n to vectors in \mathbb{R}^m .

This transformation is *linear*:

1. If A is an $m \times n$ matrix and $u, v \in \mathbb{R}^n$ then $A(u+v) = Au + Av$.

2. If A is an $m \times n$ matrix and $v \in \mathbb{R}^n$ and $c \in \mathbb{R}$ then $A(cv) = c(Av)$.

Let A and v be the general $m \times n$ matrix and n -row vector given by

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Quick way to compute Av : match up entries in the i th column of A with the entry in the i th row of v .

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix}.$$

For example, $\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix} = 1 \cdot 5 + 2 \cdot 6 + 3 \cdot 7 + 4 \cdot 8 = 5 + 12 + 21 + 32 = 70.$

3 Matrix equations

If A is an $m \times n$ matrix with columns $a_1, a_2, \dots, a_n \in \mathbb{R}^m$ and

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{R}^m$$

where each x_i is a variable, then we call $Ax = b$ a *matrix equation*.

Proposition. The matrix equation $Ax = b$ has the same solutions as both the vector equation $x_1a_1 + x_2a_2 + \cdots + x_na_n = b$ and the linear system whose augmented matrix is $\begin{bmatrix} a_1 & a_2 & \cdots & a_n & b \end{bmatrix}$.

Proposition. The matrix equation $Ax = b$ has a solution if and only if b is a linear combination of the columns of A , that is, $b \in \mathbb{R}\text{-span}\{a_1, a_2, \dots, a_n\}$.

Example. Let $A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$ and $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$.

Does $Ax = b$ have a solution for all choices of $b_1, b_2, b_3 \in \mathbb{R}$?

The system $Ax = b$ has a solution if and only if

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix}$$

is the augmented matrix of a consistent linear system. We can determine if this system is consistent by row reducing the matrix to echelon form:

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & 4b_1 + b_2 \\ 0 & 7 & 5 & 3b_1 + b_3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & 4b_1 + b_2 \\ 0 & 0 & 0 & b_1 - \frac{1}{2}b_2 + b_3 \end{bmatrix}.$$

The last matrix is in echelon form, so its leading entries are the pivot positions of our first matrix. The corresponding linear system is consistent if and only if the last column does not contain a pivot position. This occurs precisely when $b_1 - \frac{1}{2}b_2 + b_3 = 0$.

But we can choose numbers such that $b_1 - \frac{1}{2}b_2 + b_3 \neq 0$: take $b_1 = 1$ and $b_2 = b_3 = 0$. Therefore our original matrix equation $Ax = b$ does not always have a solution.

We can generalize this example:

Theorem. Let A be an $m \times n$ matrix. The following properties are equivalent, meaning that if one of them holds, then they all hold, but if one of them fails to hold, then they all fail:

1. For each vector $b \in \mathbb{R}^m$, the matrix equation $Ax = b$ has a solution.
2. Each vector $b \in \mathbb{R}^m$ is a linear combination of the columns of A .
3. The span of the columns of A is the set \mathbb{R}^m (say this as: “the columns of A span \mathbb{R}^m ”).
4. A has a pivot position in every row.

Proof. (1)-(3) are different ways of saying the same thing. Consider (4): if A has a pivot position in every row, then the augmented matrix $[A \ b]$ cannot have a pivot position in the last column. Saying that A has a pivot position in every row means that $[A \ b]$ has to be row equivalent to something like

$$\begin{bmatrix} 0 & 1 & * & * & * & c_1 \\ 0 & 0 & 0 & 4 & * & c_2 \\ 0 & 0 & 0 & 0 & 3 & c_3 \end{bmatrix}$$

where c_1, c_2, c_3 are some numbers that depend on b_1, b_2, b_3 . Regardless of what c_1, c_2, c_3 are, the given matrix has pivot columns 2, 4 and 5 but not 6.

We saw last time that not having a pivot position in the last column means that $[A \ b]$ is the augmented matrix of a consistent linear system. On the other hand, if A doesn't have a pivot position in some row, then it is always possible to choose b such that $[A \ b]$ has a pivot position in the last column, in which case the corresponding linear system has no solution. (Think about why this is true!) \square

4 Linear independence

Let v_1, v_2, \dots, v_p be vectors in \mathbb{R}^n . These vectors are *linearly independent* if the only solution to the vector equation $x_1v_1 + x_2v_2 + \dots + x_pv_p = 0$ is given by $x_1 = x_2 = \dots = x_p = 0$.

The vectors v_1, v_2, \dots, v_p are *linearly dependent* otherwise, that is, if there are numbers $c_1, c_2, \dots, c_p \in \mathbb{R}$, at least one of which is nonzero, such that $c_1v_1 + c_2v_2 + \dots + c_pv_p = 0$.

Example. If $v_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $v_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, and $v_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ then $v_1 + v_3 = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$ and $v_2 + v_3 = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$, so

$$2(v_1 + v_3) - (v_2 + v_3) = 2v_1 - v_2 + v_3 = 0.$$

Hence v_1, v_2, v_3 are linearly dependent.

It is usually not so easy to determine whether a given list of vectors is linearly independent or not.

The following result gives a way to check this for a general set of vectors:

Theorem. The columns of a matrix A are linearly independent if and only if A has a pivot position in every column.

Proof. Suppose the columns of A are $v_1, v_2, \dots, v_p \in \mathbb{R}^n$.

Saying these vectors are linearly dependent is the same thing as saying that the $n \times (p + 1)$ matrix $\begin{bmatrix} v_1 & v_2 & \dots & v_p & 0 \end{bmatrix}$ is the augmented matrix of a linear system with more than one solution (besides the trivial solution that sets all variables to zero), which happens when this system has at least one free variable (since the last column has only zeros so can never be a pivot column).

A variable x_i is free for this system precisely when i is not a pivot column of A . Thus, the columns of A are **not** linearly independent if and only if A does **not** have a pivot position in every column. \square

Theorem. Suppose $v_1, v_2, \dots, v_p \in \mathbb{R}^n$. If $p > n$ then these vectors are linearly dependent.

Proof. The $n \times p$ matrix $A = \begin{bmatrix} v_1 & v_2 & \dots & v_p \end{bmatrix}$ has at most $\min(n, p)$ pivot columns, because each column contains at most one pivot position, and each row contains at most one pivot position. Therefore if $p > n$ then A does not have a pivot position in every column so its columns are linearly dependent. \square

5 Vocabulary

Keywords from today's lecture:

1. The **product** of a matrix A and a vector v .

This is only defined if A is $m \times n$ and $v \in \mathbb{R}^n$.

In this case, if

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

then their product is

$$Av = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix} \in \mathbb{R}^m.$$

$$\text{Example: } \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \\ 7 \\ 8 \end{bmatrix} = \begin{bmatrix} 5 + 12 + 21 + 32 \\ 6 + 8 \end{bmatrix} = \begin{bmatrix} 70 \\ 14 \end{bmatrix}.$$

2. A **matrix equation**.

An equation of the form $Ax = b$ where A is an $m \times n$ matrix with columns $a_1, a_2, \dots, a_n \in \mathbb{R}^m$ and

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

is a vector where each x_i is a variable and $b \in \mathbb{R}^m$.

This equation has the same solutions as the linear system with augmented matrix $\begin{bmatrix} A & b \end{bmatrix}$.

There are several equivalent ways of characterizing whether this system has a solution.

$$\text{Example: } \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

3. **Linearly independent** vectors.

The vectors $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ are linearly independent when $x_1v_1 + \cdots + x_pv_p = 0$ if and only if $x_1 = x_2 = \cdots = x_p = 0$; equivalently, when the matrix equation

$$\begin{bmatrix} v_1 & v_2 & \cdots & v_p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = 0$$

has no solutions other than $x = 0$.

Vectors that are not linearly independent are **linearly dependent**.

Example: The three vectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$ are linearly independent.

The four vectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$, $\begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}$ are linearly dependent.