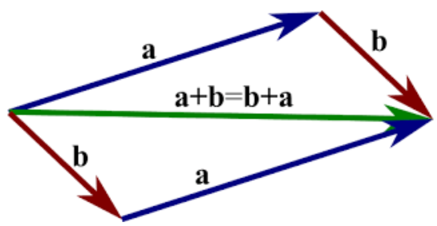
This document is a **transcript** of the lecture, with extra summary and vocabulary sections for your convenience. Due to time constraints, the notes may sometimes only contain limited illustrations, proofs, and examples; for a more thorough discussion of the course content, **consult the textbook**.

Summary

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

- A *vector* is a matrix with one column. We add and subtract vectors of the same size by doing the operations component-wise: $\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \pm \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_1 \pm v_1 \\ u_2 \pm v_2 \\ u_3 \pm v_3 \end{bmatrix}$ and $c \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix}$ for $c \in \mathbb{R}$.
- Let n be a positive integer and define $\mathbb{R}^n$ to be the set of vectors with n rows.
- We reuse the symbol 0 to mean the vector in $\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n$ whose entries are all zeros.
- Visualize vectors $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \in \mathbb{R}^2$ as arrows from the origin $(x, y) = (0, 0)$ to $(x, y) = (a_1, a_2)$.

  The sum $a + b$ for $a, b \in \mathbb{R}^2$ is then the diagonal of the parallelogram with sides a and b:



- A *linear combination* of vectors $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ is a vector of the form $c_1v_1 + c_2v_2 + \cdots + c_pv_p$ where $c_1, c_2, \dots, c_p \in \mathbb{R}$ are numbers. The set of all linear combinations of $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ is called the *span* of the vectors and is denoted by $\mathbb{R}\text{-span}\{v_1, v_2, \dots, v_p\}$.

  Example: if $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ then $\mathbb{R}\text{-span}\{e_1, e_3\} = \left\{ \begin{bmatrix} a \\ 0 \\ b \\ 0 \end{bmatrix} : a, b \in \mathbb{R} \right\}$.
- If $x_1, x_2, \dots, x_n$ are variables and $a_1, a_2, \dots, a_n, b \in \mathbb{R}^m$ are vectors then we refer to

$$x_1a_1 + x_2a_2 + \cdots + x_na_n = b$$

  as a *vector equation*. It has the same solutions as the linear system with augmented matrix

$$\begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_n & b \end{bmatrix}.$$

  The vector $b \in \mathbb{R}^m$ is in the span of the vectors $a_1, a_2, \dots, a_n \in \mathbb{R}^m$ precisely when this linear system has a solution. (And we can figure out if this happens by computing the reduced echelon form of the system's augmented matrix and checking whether the last column contains a pivot.)

1 Last time: row reduction to (reduced) echelon form

The *leading entry* in a nonzero row of a matrix is the first nonzero entry from left going right. For example, the row $\begin{bmatrix} 0 & 0 & 7 & 0 & 5 \end{bmatrix}$ has leading entry 7, which occurs in the 3rd column.

Definition. A matrix with m rows and n columns is in *echelon form* if it has the following properties:

1. If a row is nonzero, then every row above it is also nonzero.
2. The leading entry in a nonzero row is in a column to the right of the leading entry in the row above.
3. If a row is nonzero, then every entry below its leading entry in the same column is zero.

In echelon form: $\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Not in echelon form: $\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 5 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 3 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 2 & 0 & 0 \\ 1 & 3 & 5 & 0 \\ 0 & 0 & 4 & 5 \end{bmatrix}$.

Definition. A matrix is in *reduced echelon form* if

1. The matrix is in echelon form.
2. Each nonzero row has leading entry 1.
3. The leading 1 in each nonzero row is the only nonzero number in its column.

Theorem. Each matrix A is row equivalent to exactly one matrix in reduced echelon form.

We denote this matrix by $\mathsf{RREF}(A)$.

The *row reduction algorithm* is a way of constructing $\mathsf{RREF}(A)$ from A. This algorithm is something you should memorize and be able to perform quickly. The algorithm is illustrated by the following example:

Example. Writing $\to$ to indicate a sequence of row operations, we have

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 4 & 1 \\ 1 & 3 & 9 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 2 & 8 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

and the last matrix is the reduced echelon form of the first matrix.

Consider the nonzero rows of $\mathsf{RREF}(A)$. In these rows find the first nonzero entry from left to right.

If one of these leading entries is in column j, then j is a *pivot column* of A. For example if

$$\mathsf{RREF}(A) = \begin{bmatrix} 1 & -2 & 0 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

then the leading entries are in positions $(1, 1)$ and $(2, 3)$ so the pivot columns of A are 1 and 3.

If A is the augmented matrix of a linear system in variables $x_1, x_2, \dots, x_n$, then we say that x_i is a *basic variable* if i is a pivot column of A and that x_i is a *free variable* if i is not a pivot column of A.

To determine the basic and free variables of the system, we have to perform the row reduction algorithm to figure out what $\mathsf{RREF}(A)$ is first. Once we have done this, we can conclude that:

- The system has 0 solutions if the last column is a pivot column of A.
- The system has ∞ solutions if the last column is not a pivot column but there is ≥ 1 free variable.
- The system has 1 solution if there are no free variables, and the last column is not a pivot column.

Moreover, here's how you can solve the system: write down the equations in the linear system whose augmented matrix is $\mathsf{RREF}(A)$. Each nontrivial equation starts with a basic variable x_i and has the form

$$x_i + (\text{ an expression involving free variables }) = (\text{ a number })$$

After moving the expression involving free variables to the right side of the equation we get

$$x_i = (\text{ a number }) - (\text{ an expression involving free variables }).$$

To form the general solution to our original linear system, we just choose arbitrary values for the free variables and express the basic variables using these equations.

Example. The linear system

$$\begin{cases} 3x_2 - 6x_3 = 6 \\ 3x_1 - 7x_2 + 8x_3 = -5 \\ 3x_1 - 9x_2 + 12x_3 = -9 \end{cases} \qquad \text{has augmented matrix} \qquad A = \begin{bmatrix} 0 & 3 & -6 & 6 \\ 3 & -7 & 8 & -5 \\ 3 & -9 & 12 & -9 \end{bmatrix}$$

Its reduced echelon form is

$$\mathsf{RREF}(A) = \begin{bmatrix} 1 & 0 & -2 & 3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

This means that the pivot columns of A are columns 1 and 2, so x_1 and x_2 are basic variables while x_3 is a free variable. The last column is not a pivot column, so the linear system has infinitely many solutions.

The linear system with augmented matrix $\mathsf{RREF}(A)$ is

$$\begin{cases} x_1 - 2x_3 = 3 \\ x_2 - 2x_3 = 2 \\ 0 = 0 \end{cases} \quad \text{which we can rewrite as} \quad \begin{cases} x_1 = 3 + 2x_3 \\ x_2 = 2 + 2x_3 \\ 0 = 0. \end{cases}$$

We choose an arbitrary value for the free variable $x_3 = a \in \mathbb{R}$.

Then the general solution is $(x_1, x_2, x_3) = (3 + 2a, 2 + 2a, a)$ where a can be any number.

Corollary. Suppose a linear system with m equations and n variables has augmented matrix A.

If $\mathsf{RREF}(A)$ has the form $\left[\begin{array}{cccc|c} 1 & & & & b_1 \\ & 1 & & & b_2 \\ & & \ddots & & \vdots \\ & & & 1 & b_m \end{array}\right]$ where all blank entries are zero, then the linear system has exactly one solution, and this solution is given by $(x_1, x_2, \dots, x_m) = (b_1, b_2, \dots, b_m)$.

Proof. The starting linear system has the same solutions as the linear system whose augmented matrix is $\mathsf{RREF}(A)$. But the second system consists of the equations $x_1 = b_1$, $x_2 = b_2$, $\dots$, $x_m = b_m$. $\square$

2 Vectors

Until we discuss vector spaces, the term *vector* will always refer to a matrix with exactly one column:

$$\begin{bmatrix} 1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 5 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \sqrt{7} \\ \sqrt{6} \end{bmatrix}.$$

Sometimes people refer to vectors defined like this as *column vectors*.

We write a general vector as $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ where each v_i is a real number.

Two vectors u and v are equal if they have the same number of rows and the same entries in each row.

The *size* of a vector is its number of rows. We can add two vectors of the same size:

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}.$$

Note that $u + v = v + u$. If u and v don't have the same size then $u + v$ is not defined.

If v is a vector and $c \in \mathbb{R}$ is a *scalar*, i.e., a real number, then we define

$$cv = c \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}.$$

We call the new vector cv the *scalar multiple* of v by c.

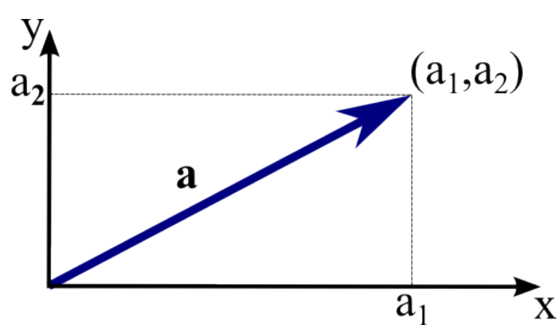
For example, we have

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} \qquad \text{and} \qquad - \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

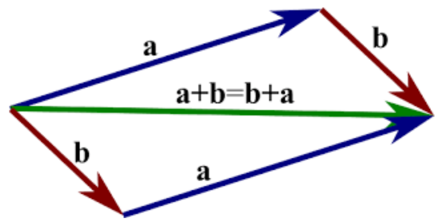
We define *subtraction* of vectors as addition after multiplying by the scalar -1:

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} - \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + (-1) \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} -1 \\ -5 \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \end{bmatrix}.$$

We write $\mathbb{R}^n$ for the set of all vectors with exactly n rows. Vectors $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \in \mathbb{R}^2$ can be identified with arrows in the Cartesian plane from the origin to the point $(x, y) = (a_1, a_2)$:



Proposition. 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 that is the opposite vertex of the parallelogram with sides a and b:



Proof. We can write $\frac{a_2}{a_1} = \frac{(a_2+b_2)-b_2}{(a_1+b_1)-b_1}$ and $\frac{b_2}{b_1} = \frac{(a_2+b_2)-a_2}{(a_1+b_1)-a_1}$.

The fractions $\frac{a_2}{a_1}$ and $\frac{b_2}{b_1}$ are the slopes of the lines through the origin containing the vectors a and b.

The other two fractions are the slopes of the lines (1) between the endpoints of b and $a+b$ and (2) between the endpoints of a and $a+b$.

The first line of the proof shows that line (1) is parallel to a, and line (2) is parallel to b.

Therefore lines (1) and (2) are the other two sides of the unique parallelogram with sides a and b.

The endpoint of $a+b$ is where lines (1) and (2) intersect.

Therefore this endpoint is the vertex of the parallelogram opposite the origin. □

The *zero vector* in $\mathbb{R}^n$ is the vector $\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ whose entries are all zero.

We use the same symbol "0" to mean both the number zero and the zero vector in $\mathbb{R}^n$ for any n. You may have to use context to figure out which number or zero vector "0" means in a given expression.

We have $0 + v = v + 0 = v$ for any vector v.

Definition. Suppose $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ are vectors and $c_1, c_2, \dots, c_p \in \mathbb{R}$ are *scalars*, i.e., numbers.

The vector $y = c_1v_1 + c_2v_2 + \cdots + c_pv_p$ is called a *linear combination* of $v_1, v_2, \dots, v_p$.

We say that y is "the linear combination of $v_1, v_2, \dots, v_p$ with *coefficients* $c_1, c_2, \dots, c_p$."

Example. Suppose $a = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$ and $b = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$ and $c = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$. Is c a linear combination of a and b?

If it were, we could find numbers $x_1, x_2 \in \mathbb{R}$ such that $x_1a + x_2b = c$, or equivalently such that

$$\begin{aligned} x_1 + 2x_2 &= 7 \\ -2x_1 + 5x_2 &= 4 \\ -5x_1 + 6x_2 &= -3. \end{aligned}$$

So to answer our question we need to determine if this linear system has a solution.

To do this, use row reduction:

$$A = \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \to \mathsf{RREF}(A) = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

The pivot columns of A are 1 and 2: the last column is **not** a pivot column. Therefore our linear system is consistent, which means that c **is** a linear combination of a and b.

We generalize this example with the following statement.

Proposition. A vector equation of the form $x_1a_1+x_2a_2+\cdots+x_na_n = b$ where $x_1, x_2, \dots, x_n$ are variables and $a_1, a_2, \dots, a_n, b \in \mathbb{R}^m$ are vectors, has the same solutions as the linear system with augmented matrix

$$\begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_n & b \end{bmatrix}. \tag{**}$$

This notation means the matrix whose ith column is a_i and last column is b.

In other words, the vector b is a linear combination of $a_1, a_2, \dots, a_n$ if and only if the linear system whose augmented matrix is (**) is consistent.

Definition. The *span* of a list of vectors $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ is the set of all vectors $y \in \mathbb{R}^n$ that are linear combinations of $v_1, v_2, \dots, v_p$. We denote this (usually infinite) set by

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

Corollary. If $v_1, v_2, \dots, v_p \in \mathbb{R}^n$, then a vector $y \in \mathbb{R}^n$ belongs to $\mathbb{R}$-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.

What does $\mathbb{R}$-span$\{v_1, v_2, \dots, v_p\}$ look like?

We can visualize the span of the 0 vector as the single point consisting of just the origin. We imagine the span of a collection of vectors that all belong to the same line through the origin as that line.

In $\mathbb{R}^2$, if the span of $v_1, v_2, \dots, v_p$ does not consist of a line, then the span is the whole plane.

To see this, imagine we have two vectors $u, v \in \mathbb{R}^2$ that are not parallel. Then we can get to any point in the plane by traveling some distance in the u direction, then some distance in the v direction. In other words, we can get any vector in $\mathbb{R}^2$ as the linear combination $au + bv$ for some scalars $a, b \in \mathbb{R}$. Draw a picture to illustrate this to yourself.

3 Vocabulary

Keywords from today's lecture:

1. **Vector**.

   A vertical list of numbers. Equivalently, a matrix with one column.

   The set of all vectors with n rows is written $\mathbb{R}^n$.

   Example: $\begin{bmatrix} 1 \\ 0 \\ -5.2 \\ 3 \end{bmatrix}$ or $\begin{bmatrix} 4 \end{bmatrix}$ or $\begin{bmatrix} \sqrt{2} \\ \pi \end{bmatrix}$.

2. **Scalar**.

   Another word for "number" or "constant." We can multiply scalars together, but not vectors.

   Example: 5 or π or $\sqrt{2}$.

3. The **zero vector** $0 \in \mathbb{R}^n$.

   The vector $\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ with n rows all equal to zero.

4. **Linear combination** of vectors.

   If $u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ are vectors, then $u + v = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}$.

   If $c \in \mathbb{R}$ is a scalar then $cv = \begin{bmatrix} cv_1 \\ cv_2 \end{bmatrix}$.

   The linear combination of vectors $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ with coefficients $a_1, a_2, \dots, a_p \in \mathbb{R}$ is

   $$a_1 v_1 + a_2 v_2 + \cdots + a_p v_p \in \mathbb{R}^n.$$

   Example: $2 \begin{bmatrix} 1 \\ 4 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \pi \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 - 0 + \pi \\ 8 - 1 + 3\pi \end{bmatrix} = \begin{bmatrix} 2 + \pi \\ 7 + 3\pi \end{bmatrix}$.

5. The **span** of a list of vectors $v_1, v_2, \dots, v_p \in \mathbb{R}^n$.

   The set of all linear combinations of the vectors $v_1, v_2, \dots, v_p \in \mathbb{R}^n$.

   A vector $u \in \mathbb{R}^n$ belongs to the span of $v_1, v_2, \dots, v_p \in \mathbb{R}^n$ if and only if the $n \times (p+1)$ matrix

   $$A = \begin{bmatrix} v_1 & v_2 & \cdots & v_p & u \end{bmatrix}$$

   is the augmented matrix of a consistent linear system.

   This happens precisely when A has no pivot positions in the last column.