Remember from last week:

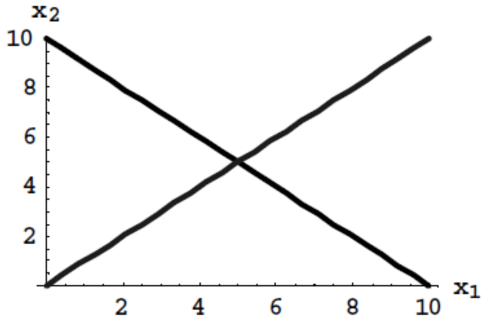
Fact: A linear system has either

- exactly one solution
- infinitely many solutions
- no solutions

We gave an algebraic proof via row reduction, but the picture, although not a proof, is useful for understanding this fact.

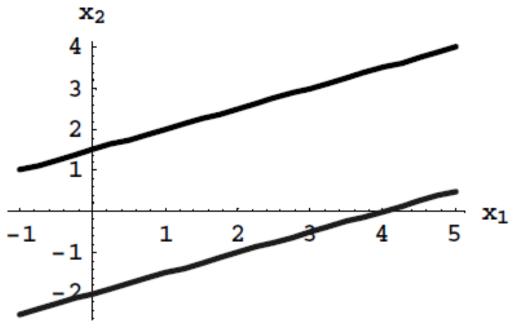
EXAMPLE Two equations in two variables:

$$\begin{aligned} x_1 + x_2 &= 10 \\ -x_1 + x_2 &= 0 \end{aligned}$$



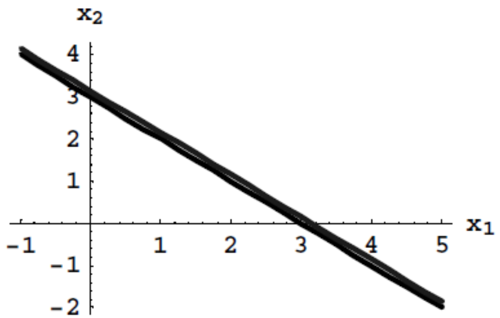
one unique solution

$$\begin{aligned} x_1 - 2x_2 &= -3 \\ 2x_1 - 4x_2 &= 8 \end{aligned}$$



no solution

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



infinitely many solutions

This week and next week, we will think more geometrically about linear systems.

§1.3-1.4 Span - related to existence of solutions
§1.5 A geometric view of solution sets (a detour)
§1.7 Linear independence - related to uniqueness of solutions

We are aiming to understand the two key concepts in three ways:

- The related computations: to solve problems about a specific linear system with numbers (Week 2 p10, Week 3 p8-9).
- The rigorous definition: to prove statements about an abstract linear system (Week 2 p15, Week 3 p11).
- The conceptual idea: to guess whether statements are true, to develop a plan for a proof or counterexample, and to help you remember the main theorems (Week 2 p13-14, Week 3 p3-5).

§1.3: Vector Equations

A column vector is a matrix with only one column.

Until Chapter 4, we will say "vector" to mean "column vector".

§1.3: Vector Equations

A column vector is a matrix with only one column.

Until Chapter 4, we will say "vector" to mean "column vector".

A vector $\mathbf{u}$ is in $\mathbb{R}^n$ if it has n rows, i.e. $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$

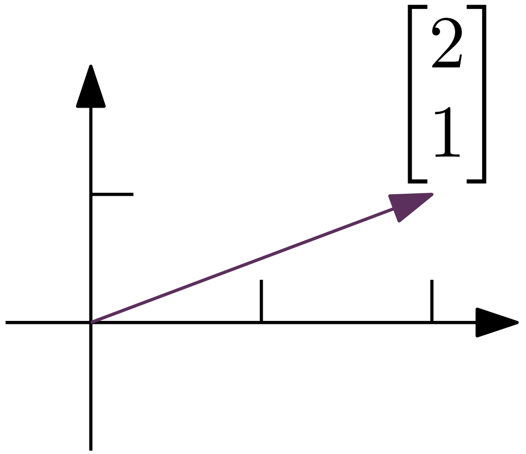
Example: $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ are vectors in $\mathbb{R}^2$.

§1.3: Vector Equations

A column vector is a matrix with only one column.

Until Chapter 4, we will say "vector" to mean "column vector".

A vector $\mathbf{u}$ is in $\mathbb{R}^n$ if it has n rows, i.e. $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$



Example: $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ are vectors in $\mathbb{R}^2$.

Vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ have a geometric meaning: think of $\begin{bmatrix} x \\ y \end{bmatrix}$ as the point (x, y) in the plane.

There are two operations we can do on vectors:

addition: if $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, then $\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$.

There are two operations we can do on vectors:

addition: if $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, then $\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$.

scalar multiplication: if $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and c is a number (a scalar), then $c\mathbf{u} = \begin{bmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{bmatrix}$.

These satisfy the usual rules for arithmetic of numbers, e.g.

$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}, \quad c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}, \quad 0\mathbf{u} = \mathbf{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Combining the operations of addition and scalar multiplication:

Definition: Given vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in $\mathbb{R}^n$ and scalars $c_1, c_2, \ldots, c_p$, the vector

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

is a *linear combination* of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ with *weights* $c_1, c_2, \ldots, c_p$.

Combining the operations of addition and scalar multiplication:

Definition: Given vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in $\mathbb{R}^n$ and scalars $c_1, c_2, \ldots, c_p$, the vector

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

is a *linear combination* of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ with *weights* $c_1, c_2, \ldots, c_p$.

Example: $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Some linear combinations of $\mathbf{u}$ and $\mathbf{v}$ are:

$$3\mathbf{u} + 2\mathbf{v} = \begin{bmatrix} 7 \\ 11 \end{bmatrix}. \qquad \tfrac{1}{3}\mathbf{u} = \begin{bmatrix} 1/3 \\ 1 \end{bmatrix}.$$

$$\mathbf{u} - 3\mathbf{v} = \begin{bmatrix} -5 \\ 0 \end{bmatrix}. \qquad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

(i.e. $\mathbf{u} + (-3)\mathbf{v}$)

Combining the operations of addition and scalar multiplication:

Definition: Given vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in $\mathbb{R}^n$ and scalars $c_1, c_2, \ldots, c_p$, the vector

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

is a *linear combination* of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ with *weights* $c_1, c_2, \ldots, c_p$.

Example: $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Some linear combinations of $\mathbf{u}$ and $\mathbf{v}$ are:

$$3\mathbf{u} + 2\mathbf{v} = \begin{bmatrix} 7 \\ 11 \end{bmatrix}. \qquad \frac{1}{3}\mathbf{u} + 0\mathbf{v} = \begin{bmatrix} 1/3 \\ 1 \end{bmatrix}.$$

$$\mathbf{u} - 3\mathbf{v} = \begin{bmatrix} -5 \\ 0 \end{bmatrix}. \qquad \mathbf{0} = 0\mathbf{u} + 0\mathbf{v} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

(i.e. $\mathbf{u} + (-3)\mathbf{v}$)

Combining the operations of addition and scalar multiplication:

Definition: Given vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ in $\mathbb{R}^n$ and scalars $c_1, c_2, \ldots, c_p$, the vector

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

is a *linear combination* of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ with *weights* $c_1, c_2, \ldots, c_p$.

Example: $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Some linear combinations of $\mathbf{u}$ and $\mathbf{v}$ are:

$$3\mathbf{u} + 2\mathbf{v} = \begin{bmatrix} 7 \\ 11 \end{bmatrix}. \qquad \frac{1}{3}\mathbf{u} + 0\mathbf{v} = \begin{bmatrix} 1/3 \\ 1 \end{bmatrix}.$$

$$\mathbf{u} - 3\mathbf{v} = \begin{bmatrix} -5 \\ 0 \end{bmatrix}. \qquad \mathbf{0} = 0\mathbf{u} + 0\mathbf{v} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

(i.e. $\mathbf{u} + (-3)\mathbf{v}$)

Study tip: an "example" after a definition does NOT mean a calculation example. These more theoretical examples are objects (vectors, in this case) that satisfy the definition, to help you understand what the definition means. You should also make your own examples when you see a definition.

Geometric interpretation of linear combinations: all the points you can go to if you are only allowed to move in the directions of $\mathbf{v}_1, \dots, \mathbf{v}_p$.

What we learned from the previous example:

1. Writing $\mathbf{b}$ as a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_p$ is the same as solving the vector equation
$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_p\mathbf{a}_p = \mathbf{b};$$
2. This vector equation has the same solution set as the linear system whose augmented matrix is
$$\left[\begin{array}{cccc|c} | & | & & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_p & \mathbf{b} \\ | & | & & | & | \end{array}\right].$$

In particular, it is not always possible to write $\mathbf{b}$ as a linear combination of given vectors: in fact, $\mathbf{b}$ is a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ if and only if there is a solution to the linear system with augmented matrix
$$\left[\begin{array}{cccc|c} | & | & & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_p & \mathbf{b} \\ | & | & & | & | \end{array}\right].$$

Definition: Suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ are in $\mathbb{R}^n$. The *span* of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$, written

$$\text{Span}\left\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\right\},$$

is the set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$.

In other words, Span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is the set of all vectors which can be written as $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p$ for any choice of weights $x_1, x_2, \ldots, x_p$.

Definition: Suppose $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ are in $\mathbb{R}^n$. The *span* of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$, written

$$\text{Span}\left\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\right\},$$

is the set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$.

In other words, Span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is the set of all vectors which can be written as $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p$ for any choice of weights $x_1, x_2, \ldots, x_p$.

In set notation:

the $\in$ sign means "is in"

$\notin$ means "is not in"

$$\text{Span}\left\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\right\} = \{\underbrace{x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p}_{} \mid \underbrace{x_1, \ldots, x_p \in \mathbb{R}}_{}\}.$$

the set of

vectors of the form $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p$

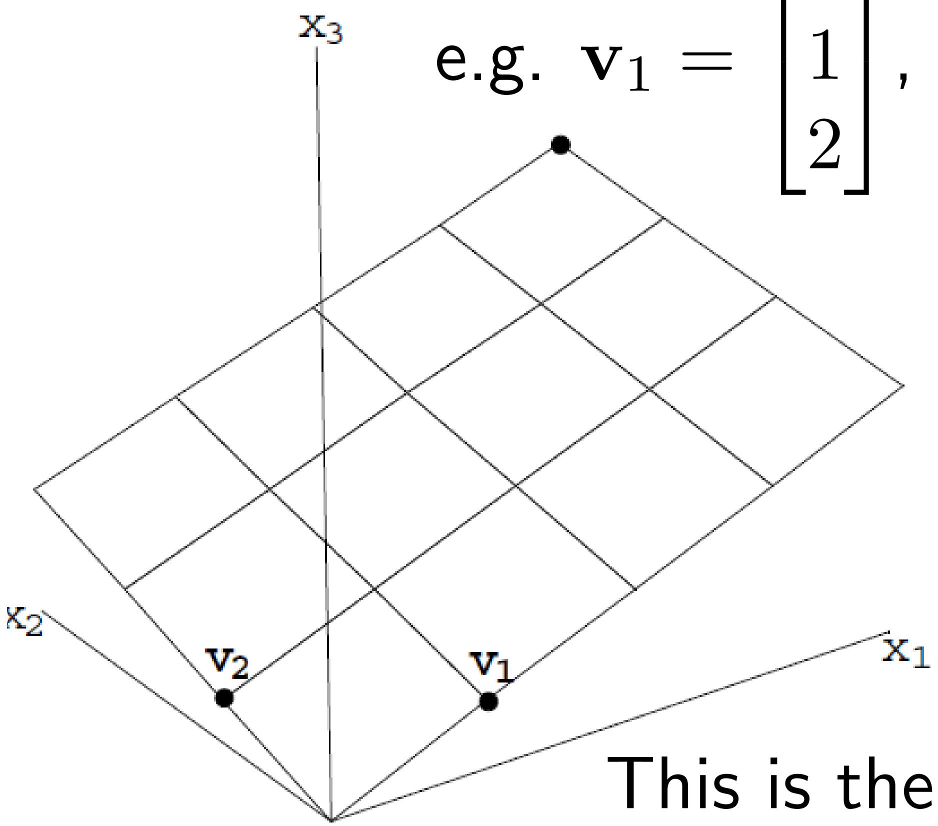
such that

$x_1, \ldots x_p$ are real numbers (i.e. they can take any value)

Example: Span of two vectors in $\mathbb{R}^3$:

When $p = 2$, the definition says Span $\{\mathbf{v}_1, \mathbf{v}_2\} = \{x_1\mathbf{v}_1 + x_2\mathbf{v}_2 \mid x_1, x_2 \in \mathbb{R}\}$.

e.g. $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$



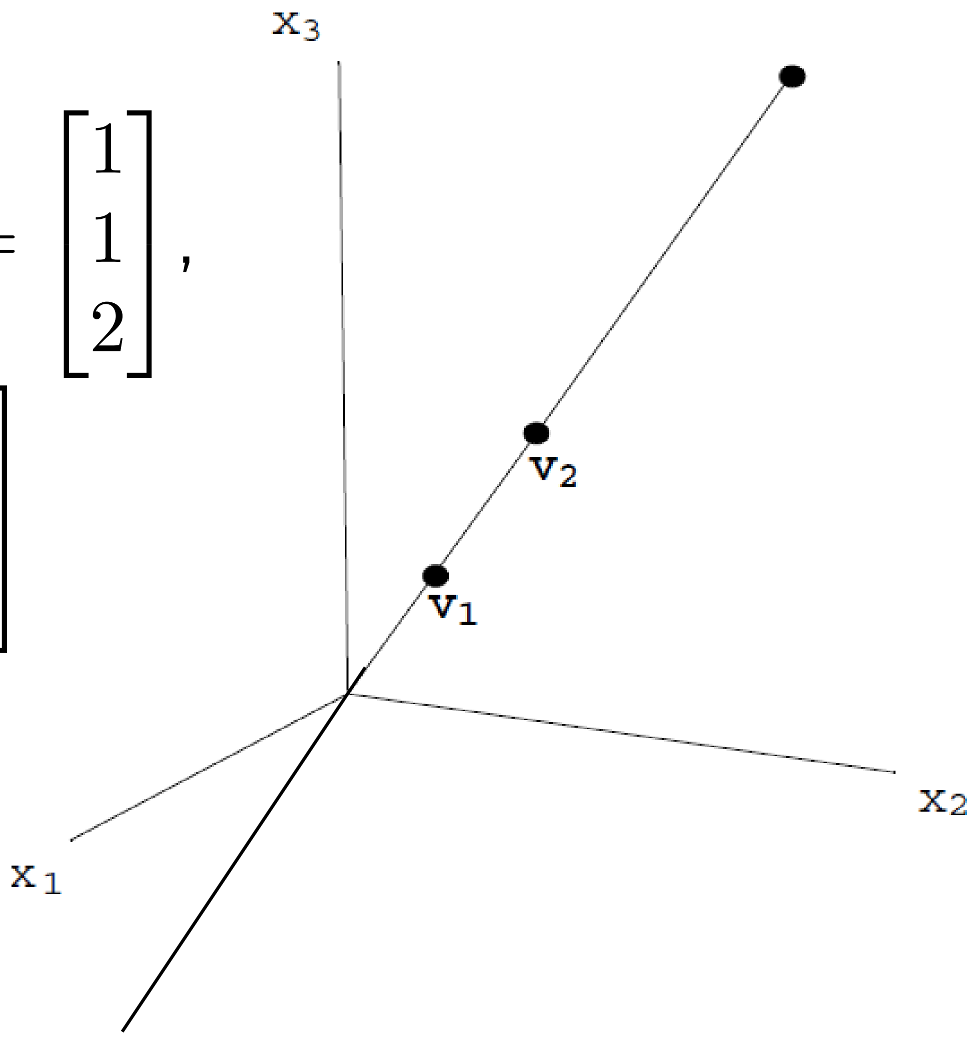
This is the plane spanned by $\{\mathbf{v}_1, \mathbf{v}_2\}$.

Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ =plane through the origin

Example: Span of two vectors in $\mathbb{R}^3$:

When $p = 2$, the definition says Span $\{\mathbf{v}_1, \mathbf{v}_2\} = \{x_1\mathbf{v}_1 + x_2\mathbf{v}_2 \mid x_1, x_2 \in \mathbb{R}\}$.

e.g. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$

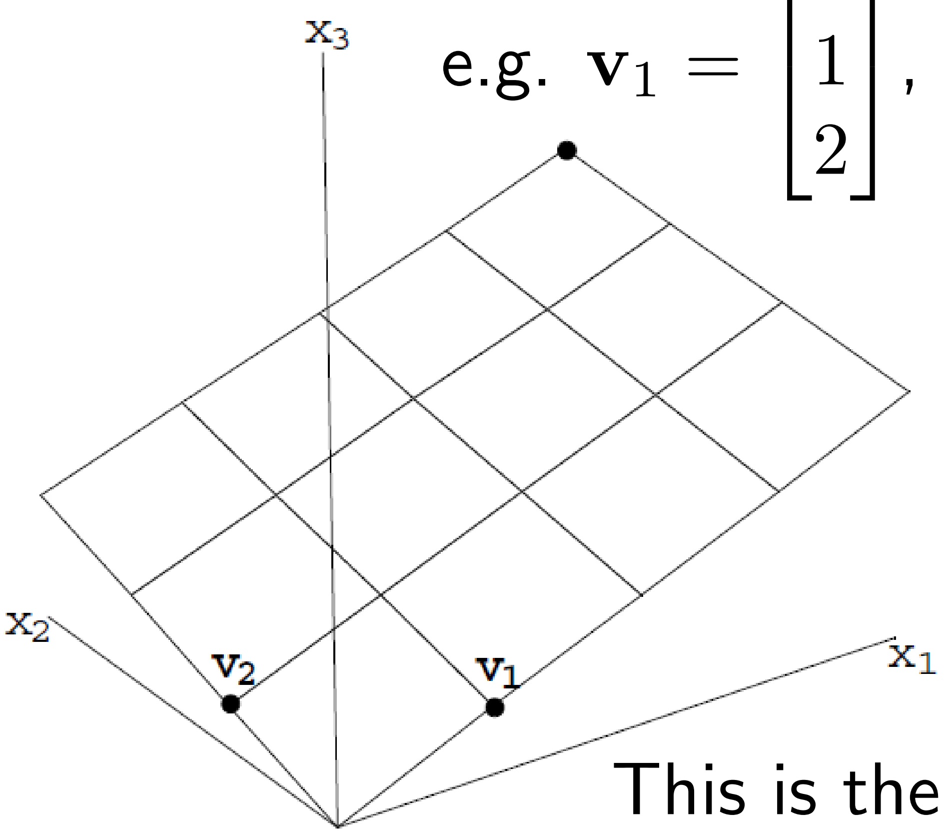


$\mathbf{v}_2$ is a multiple of $\mathbf{v}_1$

Span$\{\mathbf{v}_1, \mathbf{v}_2\}$ =**Span**$\{\mathbf{v}_1\}$ =**Span**$\{\mathbf{v}_2\}$

(line through the origin)

e.g. $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$



This is the plane spanned by $\{\mathbf{v}_1, \mathbf{v}_2\}$.

$\mathbf{v}_2$ is **not** a multiple of $\mathbf{v}_1$

Span$\{\mathbf{v}_1, \mathbf{v}_2\}$ =plane through the origin

Recall from page 10 that writing $\mathbf{b}$ as a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_p$ is equivalent to solving the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_p\mathbf{a}_p = \mathbf{b},$$

and this has the same solution set as the linear system whose augmented matrix is

$$\left[\begin{array}{cccc|c} | & | & & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_p & \mathbf{b} \\ | & | & & | & | \end{array}\right].$$

In particular, $\mathbf{b}$ is in Span $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\}$ if and only if the above linear system is consistent.

We now develop a different way to write this linear system.

§1.4: The Matrix Equation $A\mathbf{x} = \mathbf{b}$

We can think of the weights $x_1, x_2, \ldots, x_p$ as a vector.

§1.4: The Matrix Equation $A\mathbf{x} = \mathbf{b}$

We can think of the weights $x_1, x_2, \ldots, x_p$ as a vector.

m rows, p columns

The product of an $m \times p$ matrix A and a vector $\mathbf{x}$ in $\mathbb{R}^p$ is the linear combination of the columns of A using the entries of $\mathbf{x}$ as weights:

$$A\mathbf{x} = \begin{bmatrix} \vert & \vert & & \vert \\ \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_p \\ \vert & \vert & \vert & \vert \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_p\mathbf{a}_p.$$

§1.4: The Matrix Equation $A\mathbf{x} = \mathbf{b}$

We can think of the weights $x_1, x_2, \ldots, x_p$ as a vector.

m rows, p columns

The product of an $m \times p$ matrix A and a vector $\mathbf{x}$ in $\mathbb{R}^p$ is the linear combination of the columns of A using the entries of $\mathbf{x}$ as weights:

$$A\mathbf{x} = \begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_p \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_p\mathbf{a}_p.$$

Example: $\begin{bmatrix} 4 & 3 \\ 2 & 6 \\ 14 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = -2 \begin{bmatrix} 4 \\ 2 \\ 14 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 6 \\ 10 \end{bmatrix} = \begin{bmatrix} -2 \\ 8 \\ -8 \end{bmatrix}.$

Example: $\begin{bmatrix} 4 & 3 \\ 2 & 6 \\ 14 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = -2 \begin{bmatrix} 4 \\ 2 \\ 14 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 6 \\ 10 \end{bmatrix} = \begin{bmatrix} -2 \\ 8 \\ -8 \end{bmatrix}.$

There is another, faster way to compute $A\mathbf{x}$, one row of A at a time:

Example: $\begin{bmatrix} 4 & 3 \\ 2 & 6 \\ 14 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4(-2) + 3(2) \\ \\ \\ \end{bmatrix} = \begin{bmatrix} -2 \\ \\ \\ \end{bmatrix}.$

Example: $\begin{bmatrix} 4 & 3 \\ 2 & 6 \\ 14 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = -2 \begin{bmatrix} 4 \\ 2 \\ 14 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 6 \\ 10 \end{bmatrix} = \begin{bmatrix} -2 \\ 8 \\ -8 \end{bmatrix}.$

There is another, faster way to compute $A\mathbf{x}$, one row of A at a time:

Example: $\begin{bmatrix} 4 & 3 \\ 2 & 6 \\ 14 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4(-2) + 3(2) \\ 2(-2) + 6(2) \\ \\ \end{bmatrix} = \begin{bmatrix} -2 \\ 8 \\ \\ \end{bmatrix}.$

Example: $$\begin{bmatrix} 4 & 3 \\ 2 & 6 \\ 14 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = -2 \begin{bmatrix} 4 \\ 2 \\ 14 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 6 \\ 10 \end{bmatrix} = \begin{bmatrix} -2 \\ 8 \\ -8 \end{bmatrix}.$$

There is another, faster way to compute $A\mathbf{x}$, one row of A at a time:

Example: $$\begin{bmatrix} 4 & 3 \\ 2 & 6 \\ 14 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4(-2) + 3(2) \\ 2(-2) + 6(2) \\ 14(-2) + 10(2) \end{bmatrix} = \begin{bmatrix} -2 \\ 8 \\ -8 \end{bmatrix}.$$

Example: $\begin{bmatrix} 4 & 3 \\ 2 & 6 \\ 14 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = -2 \begin{bmatrix} 4 \\ 2 \\ 14 \end{bmatrix} + 2 \begin{bmatrix} 3 \\ 6 \\ 10 \end{bmatrix} = \begin{bmatrix} -2 \\ 8 \\ -8 \end{bmatrix}.$

There is another, faster way to compute $A\mathbf{x}$, one row of A at a time:

Example: $\begin{bmatrix} 4 & 3 \\ 2 & 6 \\ 14 & 10 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4(-2) + 3(2) \\ 2(-2) + 6(2) \\ 14(-2) + 10(2) \end{bmatrix} = \begin{bmatrix} -2 \\ 8 \\ -8 \end{bmatrix}.$

It is easy to check that $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$ and $A(c\mathbf{u}) = cA\mathbf{u}$.

Warning: The product $A\mathbf{x}$ is only defined if the number of columns of A equals the number of rows of x. The number of rows of $A\mathbf{x}$ is the number of rows of A.

Warning: Always write $A\mathbf{x}$, with the matrix on the left and the vector on the right - $\mathbf{x}A$ has a different meaning.

We have three ways of viewing the same problem:

1. The system of linear equations with augmented matrix $[A|\mathbf{b}]$,
2. The vector equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_p\mathbf{a}_p = \mathbf{b}$,
3. The matrix equation $A\mathbf{x} = \mathbf{b}$.

We have three ways of viewing the same problem:

1. The system of linear equations with augmented matrix $[A|\mathbf{b}]$,
2. The vector equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_p\mathbf{a}_p = \mathbf{b}$,
3. The matrix equation $A\mathbf{x} = \mathbf{b}$.

These three problems have the same solution set, so the following three things are the same (they are simply different ways to say "the above problem has a solution"):

1. The system of linear equations with augmented matrix $[A|\mathbf{b}]$ has a solution,
2. $\mathbf{b}$ is a linear combination of the columns of A (or $\mathbf{b}$ is in the span of the columns of A),
3. The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution.

We have three ways of viewing the same problem:

1. The system of linear equations with augmented matrix $[A|\mathbf{b}]$,
2. The vector equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_p\mathbf{a}_p = \mathbf{b}$,
3. The matrix equation $A\mathbf{x} = \mathbf{b}$.

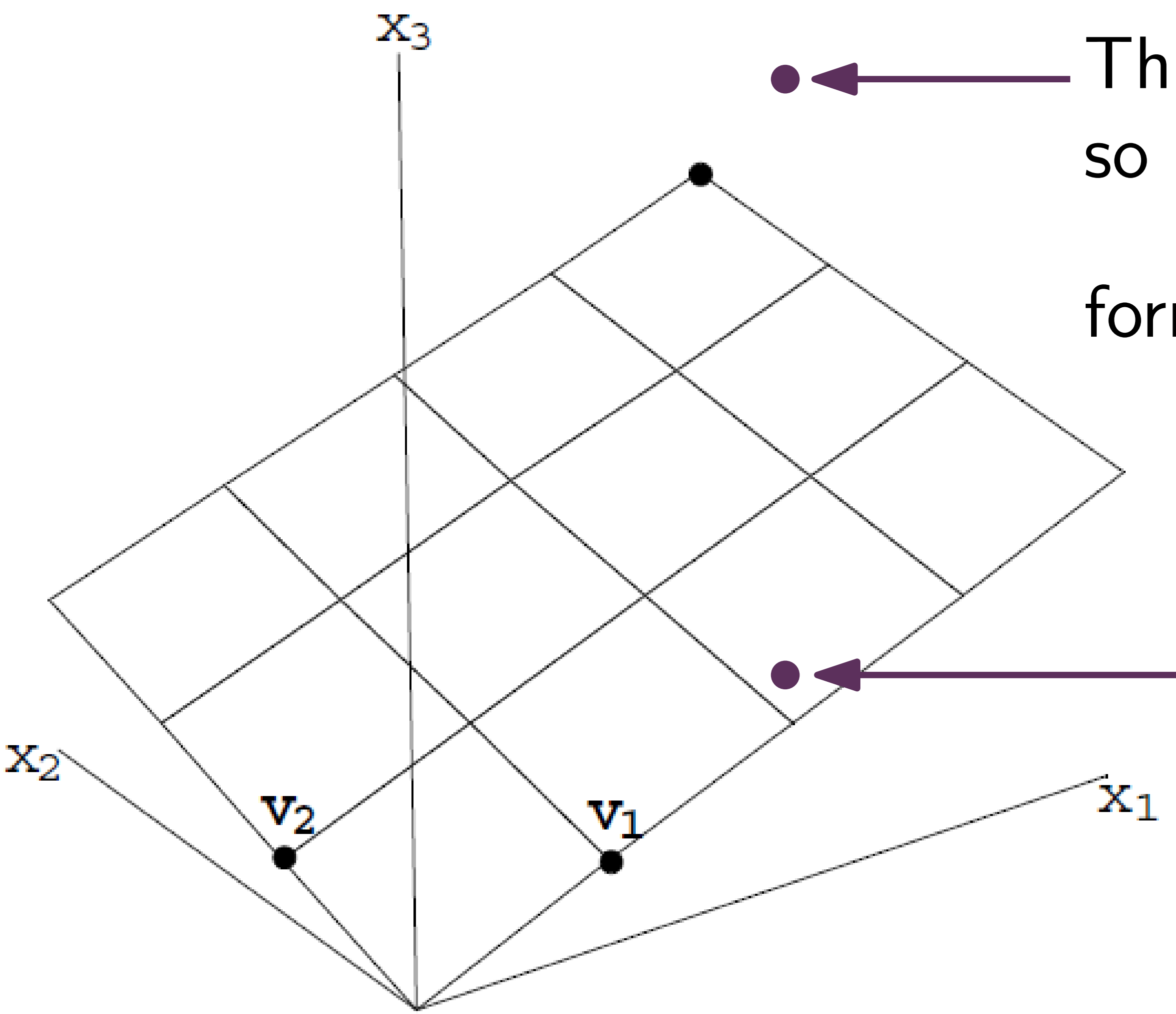
These three problems have the same solution set, so the following three things are the same (they are simply different ways to say "the above problem has a solution"):

1. The system of linear equations with augmented matrix $[A|\mathbf{b}]$ has a solution,
2. $\mathbf{b}$ is a linear combination of the columns of A (or $\mathbf{b}$ is in the span of the columns of A),
3. The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution.

Another way of saying this: The span of the columns of A is the set of vectors $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ has a solution.

The span of the columns of A is the set of vectors $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ has a solution.

Example: If $A = \begin{bmatrix} 3 & 1 \\ 1 & 4 \\ 2 & 1 \end{bmatrix}$, then the relevant vectors are $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$.



This $\mathbf{b}$ is not on the plane spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$, so $A\mathbf{x} = \mathbf{b}$ does not have a solution. The echelon form of $[A|\mathbf{b}]$ is $\left[\begin{array}{cc|c} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & a \end{array}\right]$ where $a \neq 0$.

This $\mathbf{b}$ is on the plane spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$, so $A\mathbf{x} = \mathbf{b}$ has a solution. The echelon form of $[A|\mathbf{b}]$ is $\left[\begin{array}{cc|c} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & 0 \end{array}\right]$.

Warning: If A is an $m \times n$ matrix, then the pictures on the previous page are for the right hand side $\mathbf{b} \in \mathbb{R}^m$, not for the solution $\mathbf{x} \in \mathbb{R}^n$ (as we were drawing in Week 1, and also in p28-30 later this week). In this example, we cannot draw the solution sets on the same picture, because the solutions $\mathbf{x}$ are in $\mathbb{R}^2$, but our picture is in $\mathbb{R}^3$.

So these three things are the same:

1. The system of linear equations with augmented matrix $[A|\mathbf{b}]$ has a solution,
2. $\mathbf{b}$ is a linear combination of the columns of A (or $\mathbf{b}$ is in the span of the columns of A),
3. The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution.

One question of particular interest: when are the above statements true for all vectors $\mathbf{b}$ in $\mathbb{R}^m$? i.e. when is $A\mathbf{x} = \mathbf{b}$ consistent for all right hand sides $\mathbf{b}$, and when is Span $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\} = \mathbb{R}^m$?

Example: $(m = 3)$ Let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Then Span $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \mathbb{R}^3$, because $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

But for a more complicated set of vectors, the weights will be more complicated functions of x, y, z. So we want a better way to answer this question.

Theorem 4: Existence of solutions to linear systems: For an $m \times n$ matrix A, the following statements are logically equivalent (i.e. for any particular matrix A, they are all true or all false):

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.

b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.

c.

d.

Theorem 4: Existence of solutions to linear systems: For an $m \times n$ matrix A, the following statements are logically equivalent (i.e. for any particular matrix A, they are all true or all false):

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.

b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.

c. The columns of A span $\mathbb{R}^m$ (i.e. Span $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\} = \mathbb{R}^m$).

d.

Theorem 4: Existence of solutions to linear systems: For an $m \times n$ matrix A, the following statements are logically equivalent (i.e. for any particular matrix A, they are all true or all false):

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
c. The columns of A span $\mathbb{R}^m$ (i.e. Span $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\} = \mathbb{R}^m$).
d. rref(A) has a pivot in every row.

Theorem 4: Existence of solutions to linear systems: For an $m \times n$ matrix A, the following statements are logically equivalent (i.e. for any particular matrix A, they are all true or all false):

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
c. The columns of A span $\mathbb{R}^m$ (i.e. Span $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\} = \mathbb{R}^m$).
d. rref(A) has a pivot in every row.

Warning: the theorem says nothing about the uniqueness of the solution.

Theorem 4: Existence of solutions to linear systems: For an $m \times n$ matrix A, the following statements are logically equivalent (i.e. for any particular matrix A, they are all true or all false):

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
c. The columns of A span $\mathbb{R}^m$ (i.e. Span $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\} = \mathbb{R}^m$).
d. rref(A) has a pivot in every row.

Warning: the theorem says nothing about the uniqueness of the solution.

Proof: (outline): By the previous discussion, (a), (b) and (c) are logically equivalent. So, to finish the proof, we only need to show that (a) and (d) are logically equivalent, i.e. we need to show that,

- if (d) is true, then (a) is true;
- if (d) is false, then (a) is false.

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
d. rref(A) has a pivot in every row.

Proof: (continued)

Suppose (d) is true.

So (a) is true.

Suppose (d) is false.

So (a) is false

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
d. rref(A) has a pivot in every row.

Proof: (continued)

Suppose (d) is true. Then, for every $\mathbf{b}$ in $\mathbb{R}^m$, the augmented matrix $[A|\mathbf{b}]$ row-reduces to $[\text{rref}(A)|\mathbf{d}]$ for some $\mathbf{d}$ in $\mathbb{R}^m$. This does not have a row of the form $[0 \dots 0|*]$, so, by the Existence of Solutions Theorem (Week 1 p27), $A\mathbf{x} = \mathbf{b}$ is consistent. So (a) is true.

Suppose (d) is false.

So (a) is false

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
d. rref(A) has a pivot in every row.

Proof: (continued)

Suppose (d) is true. Then, for every $\mathbf{b}$ in $\mathbb{R}^m$, the augmented matrix $[A|\mathbf{b}]$ row-reduces to $[\text{rref}(A)|\mathbf{d}]$ for some $\mathbf{d}$ in $\mathbb{R}^m$. This does not have a row of the form $[0 \dots 0|*]$, so, by the Existence of Solutions Theorem (Week 1 p27), $A\mathbf{x} = \mathbf{b}$ is consistent. So (a) is true.

Suppose (d) is false. We want to find a counterexample to (a): i.e. we want to find a vector $\mathbf{b}$ in $\mathbb{R}^m$ such that $A\mathbf{x} = \mathbf{b}$ has no solution.

(This last part of the proof, written on the next page, is hard, and is not something you are expected to think of by yourself. But you should try to understand the part of the proof on this page.)

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
d. rref(A) has a pivot in every row.

Proof: (continued) Suppose (d) is false. We want to find a counterexample to (a): i.e. we want to find a vector $\mathbf{b}$ in $\mathbb{R}^m$ such that $A\mathbf{x} = \mathbf{b}$ has no solution.

rref(A) does not have a pivot in every row, so its last row is $[0 \dots 0]$.

Example: $$\begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix} \xrightarrow{R_2 \to R_2 + 2R_1} \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}$$

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
d. rref(A) has a pivot in every row.

Proof: (continued) Suppose (d) is false. We want to find a counterexample to (a): i.e. we want to find a vector $\mathbf{b}$ in $\mathbb{R}^m$ such that $A\mathbf{x} = \mathbf{b}$ has no solution.

rref(A) does not have a pivot in every row, so its last row is $[0 \dots 0]$.

Let $\mathbf{d} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$. Then the linear system with augmented matrix $[\text{rref}(A)|\mathbf{d}]$ is inconsistent.

Example: $\begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix} \xrightarrow{R_2 \to R_2 + 2R_1} \left[\begin{array}{cc|c} 1 & -3 & 1 \\ 0 & 0 & 1 \end{array}\right]$

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
d. rref(A) has a pivot in every row.

Proof: (continued) Suppose (d) is false. We want to find a counterexample to (a): i.e. we want to find a vector $\mathbf{b}$ in $\mathbb{R}^m$ such that $A\mathbf{x} = \mathbf{b}$ has no solution.

rref(A) does not have a pivot in every row, so its last row is $[0 \dots 0]$.

Let $\mathbf{d} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$.

Then the linear system with augmented matrix $[\text{rref}(A)|\mathbf{d}]$ is inconsistent.

Now we apply the row operations in reverse to get an equivalent linear system $[A|\mathbf{b}]$ that is inconsistent.

Example:

$$\left[\begin{array}{cc|c} 1 & -3 & 1 \\ -2 & 6 & -1 \end{array}\right] \underset{R_2 \to R_2 - 2R_1}{\overset{R_2 \to R_2 + 2R_1}{\rightleftharpoons}} \left[\begin{array}{cc|c} 1 & -3 & 1 \\ 0 & 0 & 1 \end{array}\right]$$

Theorem 4: Existence of solutions to linear systems: For an $m \times n$ matrix A, the following statements are logically equivalent (i.e. for any particular matrix A, they are all true or all false):

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
c. The columns of A span $\mathbb{R}^m$ (i.e. Span $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\} = \mathbb{R}^m$).
d. rref(A) has a pivot in every row.

We will add more statements to this theorem throughout the course.

Theorem 4: Existence of solutions to linear systems: For an $m \times n$ matrix A, the following statements are logically equivalent (i.e. for any particular matrix A, they are all true or all false):

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
c. The columns of A span $\mathbb{R}^m$ (i.e. Span $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\} = \mathbb{R}^m$).
d. rref(A) has a pivot in every row.

We will add more statements to this theorem throughout the course.

Observe that rref(A) has at most one pivot per column (condition 5 of a reduced echelon form, or think about how we perform row-reduction). So if A has more rows than columns (a "tall" matrix), then rref(A) cannot have a pivot in every row, so the statements above are all false.
In particular, a set of fewer than m vectors cannot span $\mathbb{R}^m$.

Warning/Exercise: It is not true that any set of m or more vectors span $\mathbb{R}^m$: can you think of an example?

§1.5: Solution Sets of Linear Systems

Goal: use vector notation to give geometric descriptions of solution sets

to compare the solution sets of $A\mathbf{x} = \mathbf{b}$ and of $A\mathbf{x} = \mathbf{0}$.

Definition: A linear system is *homogeneous* if the right hand side is the zero vector, i.e.

$$A\mathbf{x} = \mathbf{0}.$$

§1.5: Solution Sets of Linear Systems

Goal: use vector notation to give geometric descriptions of solution sets

to compare the solution sets of $A\mathbf{x} = \mathbf{b}$ and of $A\mathbf{x} = \mathbf{0}$.

Definition: A linear system is *homogeneous* if the right hand side is the zero vector, i.e.

$$A\mathbf{x} = \mathbf{0}.$$

When we row-reduce $[A|\mathbf{0}]$, the right hand side stays $\mathbf{0}$, so the reduced echelon form does not have a row of the form $[0 \dots 0|*]$ with $* \neq 0$.
So a homogeneous system is always consistent.

§1.5: Solution Sets of Linear Systems

Goal: use vector notation to give geometric descriptions of solution sets
to compare the solution sets of $A\mathbf{x} = \mathbf{b}$ and of $A\mathbf{x} = \mathbf{0}$.

Definition: A linear system is *homogeneous* if the right hand side is the zero vector, i.e.

$$A\mathbf{x} = \mathbf{0}.$$

When we row-reduce $[A|\mathbf{0}]$, the right hand side stays $\mathbf{0}$, so the reduced echelon form does not have a row of the form $[0 \dots 0|*]$ with $* \neq 0$.
So a homogeneous system is always consistent.

In fact, $\mathbf{x} = \mathbf{0}$ is always a solution, because $A\mathbf{0} = \mathbf{0}$. The solution $\mathbf{x} = \mathbf{0}$ called the trivial solution.

A non-trivial solution $\mathbf{x}$ is a solution where at least one x_i is non-zero.

In our first example:

- The solution set of $A\mathbf{x} = \mathbf{0}$ is a line through the origin parallel to $\mathbf{v}$.
- The solution set of $A\mathbf{x} = \mathbf{b}$ is a line through $\mathbf{p}$ parallel to $\mathbf{v}$.

In our second example:

- The solution set of $A\mathbf{x} = \mathbf{0}$ is a plane through the origin parallel to $\mathbf{u}$ and $\mathbf{v}$.
- The solution set of $A\mathbf{x} = \mathbf{b}$ is a plane through $\mathbf{p}$ parallel to $\mathbf{u}$ and $\mathbf{v}$.

In both cases: to get the solution set of $A\mathbf{x} = \mathbf{b}$, start with the solution set of $A\mathbf{x} = \mathbf{0}$ and translate it by $\mathbf{p}$.

$\mathbf{p}$ is called a particular solution (one solution out of many).

In our first example:

- The solution set of $A\mathbf{x} = \mathbf{0}$ is a line through the origin parallel to $\mathbf{v}$.
- The solution set of $A\mathbf{x} = \mathbf{b}$ is a line through $\mathbf{p}$ parallel to $\mathbf{v}$.

In our second example:

- The solution set of $A\mathbf{x} = \mathbf{0}$ is a plane through the origin parallel to $\mathbf{u}$ and $\mathbf{v}$.
- The solution set of $A\mathbf{x} = \mathbf{b}$ is a plane through $\mathbf{p}$ parallel to $\mathbf{u}$ and $\mathbf{v}$.

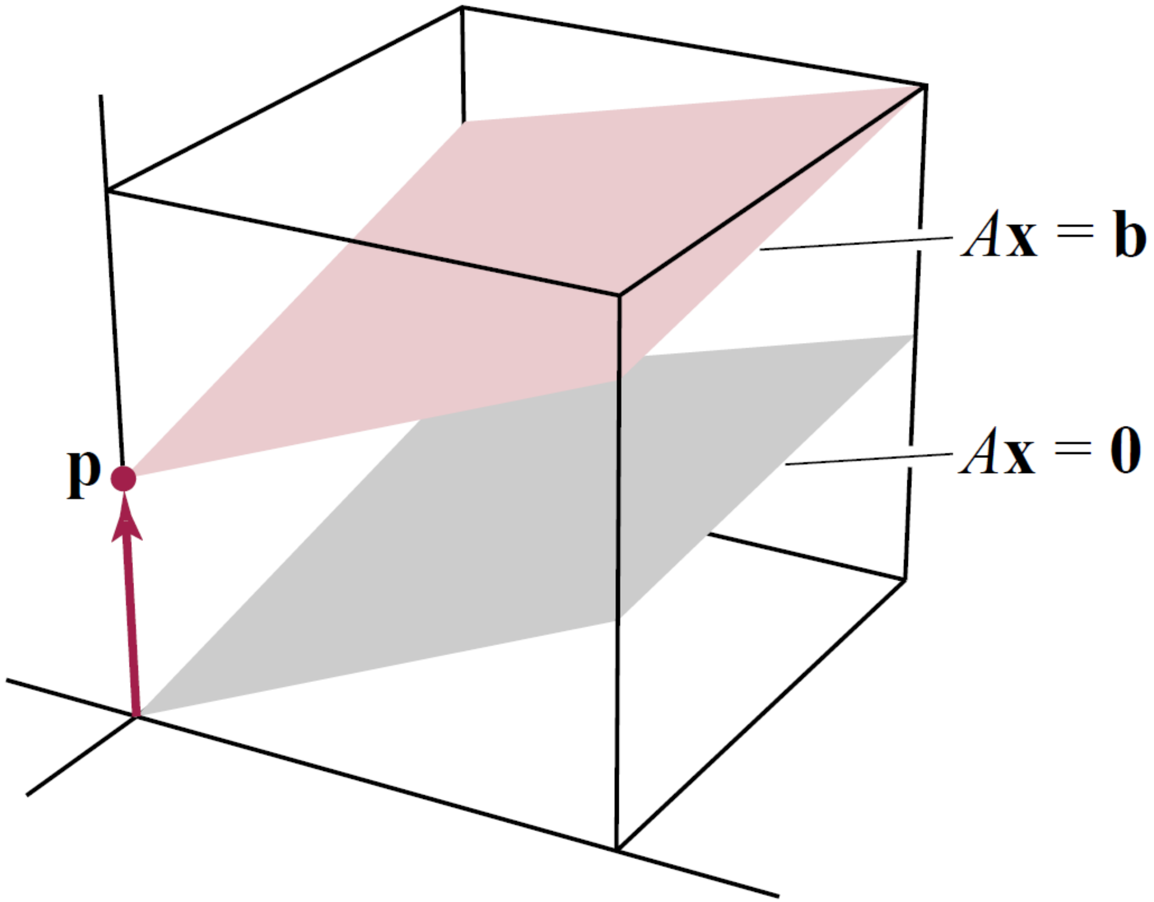
In both cases: to get the solution set of $A\mathbf{x} = \mathbf{b}$, start with the solution set of $A\mathbf{x} = \mathbf{0}$ and translate it by $\mathbf{p}$.

$\mathbf{p}$ is called a particular solution (one solution out of many).

In general:

Theorem 6: Solutions and homogeneous equations: Suppose $\mathbf{p}$ is a solution to $A\mathbf{x} = \mathbf{b}$. Then the solution set to $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v_h}$, where $\mathbf{v_h}$ is any solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$.

Theorem 6: Solutions and homogeneous equations: Suppose $\mathbf{p}$ is a solution to $A\mathbf{x} = \mathbf{b}$. Then the solution set to $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v_h}$, where $\mathbf{v_h}$ is any solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$.



Parallel solution sets of $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{x} = \mathbf{0}$.

Theorem 6: Solutions and homogeneous equations: Suppose $\mathbf{p}$ is a solution to $A\mathbf{x} = \mathbf{b}$. Then the solution set to $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v_h}$, where $\mathbf{v_h}$ is any solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$.

Proof: (outline)
We show that $\mathbf{w} = \mathbf{p} + \mathbf{v_h}$ is a solution:

$$\begin{aligned} & A(\mathbf{p} + \mathbf{v_h}) \\ =& A\mathbf{p} + A\mathbf{v_h} \\ =& \mathbf{b} + \mathbf{0} \\ =& \mathbf{b}. \end{aligned}$$

We also need to show that all solutions are of the form $\mathbf{w} = \mathbf{p} + \mathbf{v_h}$ - see q25 in Section 1.5 of the textbook.

How this theorem is useful: a shortcut to Q1b on ex. sheet #5:

Example: Let $A = \begin{bmatrix} | & | & | & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \\ | & | & | & | \end{bmatrix} = \begin{bmatrix} 1 & 3 & 0 & -4 \\ 2 & 6 & 0 & -8 \end{bmatrix}$.

In Q1a, you found that the solution set to $A\mathbf{x} = \mathbf{0}$ is $\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} r + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} s + \begin{bmatrix} 4 \\ 0 \\ 0 \\ 1 \end{bmatrix} t$, where r, s, t can take any value.

In Q1b, you want to solve $A\mathbf{x} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$. Now $\begin{bmatrix} 3 \\ 6 \end{bmatrix} = 0\mathbf{a}_1 + 1\mathbf{a}_2 + 0\mathbf{a}_3 + 0\mathbf{a}_4 = A \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, so $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ is a particular solution. So the solution set is $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} r + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} s + \begin{bmatrix} 4 \\ 0 \\ 0 \\ 1 \end{bmatrix} t$, where r, s, t can take any value.

Notice that this solution looks different from the solution obtained from row-reduction:

rref $\left(\left[\begin{array}{cccc|c} 1 & 3 & 0 & -4 & 3 \\ 2 & 6 & 0 & -8 & 6 \end{array}\right]\right) = \left[\begin{array}{cccc|c} 1 & 3 & 0 & -4 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$, which gives a different particular solution $\begin{bmatrix} 3 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

But the solution sets are the same:

$$\begin{aligned} \begin{bmatrix} 3 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} r + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} s + \begin{bmatrix} 4 \\ 0 \\ 0 \\ 1 \end{bmatrix} t &= \begin{bmatrix} 3 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} (r-1) + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} s + \begin{bmatrix} 4 \\ 0 \\ 0 \\ 1 \end{bmatrix} t \\ &= \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} (r-1) + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} s + \begin{bmatrix} 4 \\ 0 \\ 0 \\ 1 \end{bmatrix} t, \end{aligned}$$

and r, s, t taking any value is equivalent to $r-1, s, t$ taking any value.