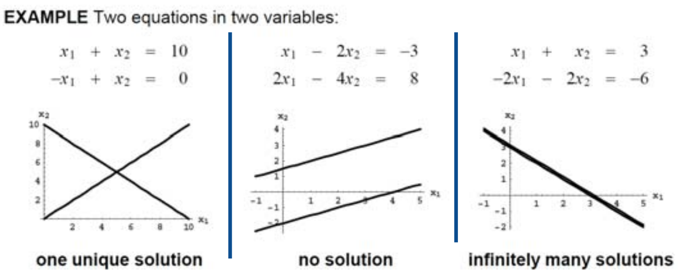
Remember from last week:

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.



Now 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:

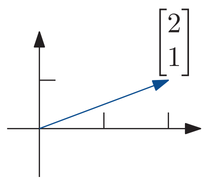
- The related computations: to solve problems about a specific linear system with numbers (p13-14, p37-38).
- The rigorous definition: to prove statements about an abstract linear system (p39-40).
- The conceptual idea: to guess whether statements are true, to develop a plan for a proof or counterexample, and to help you remember the main theorems (p11-12, p33-34).

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

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

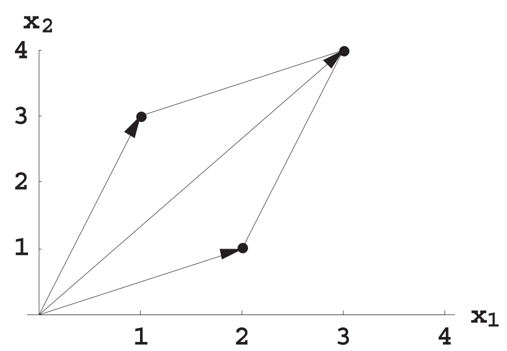
These satisfy the usual rules for arithmetic of numbers, e.g.

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

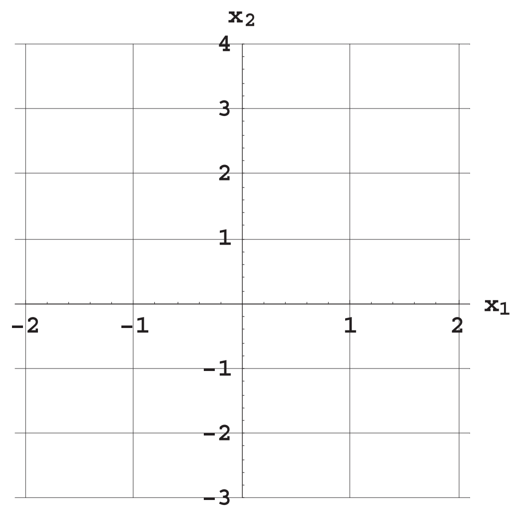
Parallelogram rule for addition of two vectors:

If $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{R}^2$ are represented as points in the plane, then $\mathbf{u}+\mathbf{v}$ corresponds to the fourth vertex of the parallelogram whose other vertices are $\mathbf{0}$, $\mathbf{u}$ and $\mathbf{v}$. (Note that $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.)

EXAMPLE: Let $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$



EXAMPLE: Let $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Express $\mathbf{u}$, $2\mathbf{u}$, and $\frac{-3}{2}\mathbf{u}$ on a graph.



Combining the operations of addition and scalar multiplication:
Definition: Given vectors $\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_p$ in $\mathbb{R}^n$ and scalars $c_1, c_2, \dots c_p$, the vector

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

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

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:

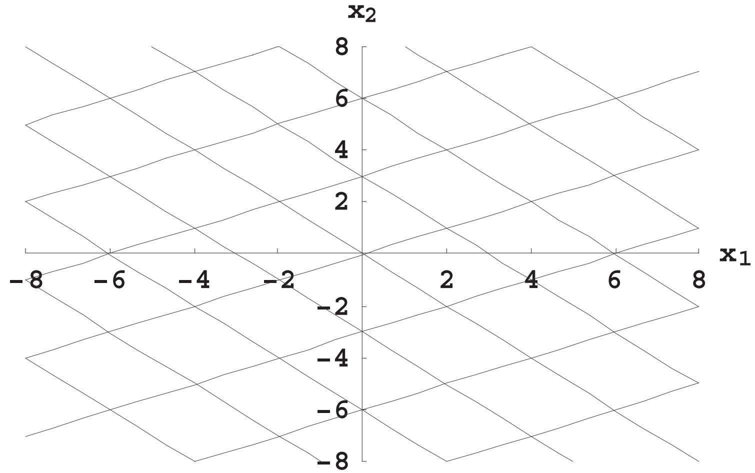
$$3\mathbf{u} + 2\mathbf{v} = \begin{bmatrix}7\\11\end{bmatrix}. \qquad \tfrac{1}{3}\mathbf{u} + 0\mathbf{v} = \begin{bmatrix}1/3\\1\end{bmatrix}.$$

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

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

EXAMPLE: Let $\mathbf{v}_1 = \begin{bmatrix}2\\1\end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix}-2\\2\end{bmatrix}$. Express each of the following as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$:

$$\mathbf{a} = \begin{bmatrix}0\\3\end{bmatrix}, \mathbf{b} = \begin{bmatrix}-4\\1\end{bmatrix}, \mathbf{c} = \begin{bmatrix}6\\6\end{bmatrix}, \mathbf{d} = \begin{bmatrix}7\\-4\end{bmatrix}$$



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

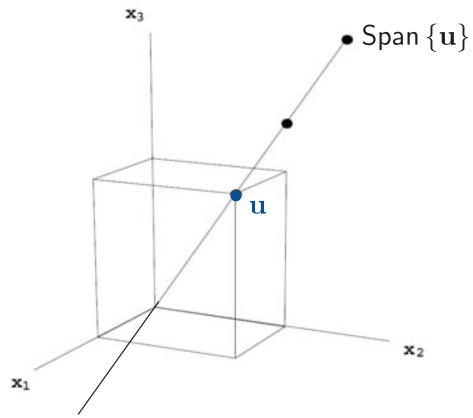
$$\text{Span}\,\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}\,,$$

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 $x_1, x_2, \ldots, x_p$.

Example: Span of one vector in $\mathbb{R}^3$

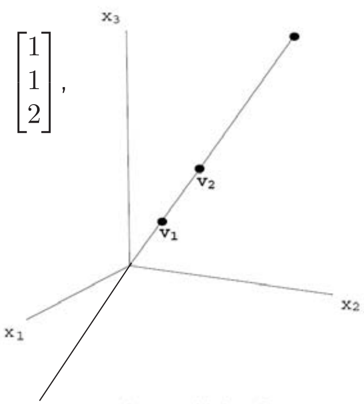
- Span $\{\mathbf{0}\} = \{\mathbf{0}\}$, because $c\mathbf{0} = \mathbf{0}$ for all scalars c.
- If $\mathbf{u}$ is not the zero vector, then Span $\{\mathbf{u}\}$ is a line through the origin in the direction $\mathbf{u}$.



We can also say "$\{\mathbf{u}\}$ spans a line through the origin".

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

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

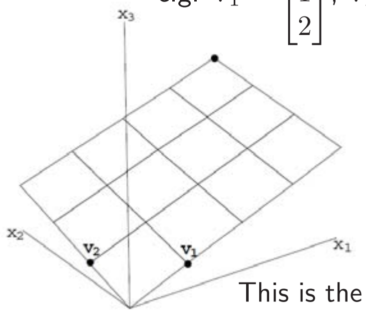


$\mathbf{v}_2$ is a multiple of $\mathbf{v}_1$

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

EXAMPLE: Let $\mathbf{a}_1 = \begin{bmatrix} 4 \\ 2 \\ 14 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 3 \\ 6 \\ 10 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} -2 \\ 8 \\ -8 \end{bmatrix}$.

Determine if **b** is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$.

Solution: Vector **b** is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$ if we can find weights x_1, x_2 such that

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}.$$

Vector equation:

Corresponding linear system:

Corresponding augmented matrix:

$$\left[\begin{array}{cc|c} 4 & 3 & -2 \\ 2 & 6 & 8 \\ 14 & 10 & -8 \end{array}\right]$$

Reduced echelon form:

$$\left[\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right]$$

From the previous example, we see that the vector equation

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

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 a linear combination of $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ (i.e. $\mathbf{b}$ is in Span $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\}$) if and only if there is a solution to the linear system with augmented matrix

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

We now develop a different way to write this equation.

§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 way to compute $A\mathbf{x}$, one row of A at a time:

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

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.

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

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

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.

(In fact, the three problems have the same solution set.)

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.

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. 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 previous the 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. Then, for every $\mathbf{b}$ in $\mathbb{R}^m$, the augmented matrix $[A|\mathbf{b}]$ row-reduces to [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 p 25), $A\mathbf{x} = \mathbf{b}$ 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.

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

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 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p\} = \mathbb{R}^m$).
d. rref(A) has a pivot in every row.

Observe that rref(A) has at most one pivot per column (condition 5 of a reduced echelon form). 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 $\mathbb{R}^m$.

§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.

If there are non-trivial solutions, what does the solution set look like?

EXAMPLE:

$$2x_1 + 4x_2 - 6x_3 = 0$$

$$4x_1 + 8x_2 - 10x_3 = 0$$

Corresponding augmented matrix:

$$\left[\begin{array}{ccc|c} 2 & 4 & -6 & 0 \\ 4 & 8 & -10 & 0 \end{array}\right]$$

Corresponding reduced echelon form:

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

Solution set:

Geometric representation:

EXAMPLE: (same left hand side as before)

$$2x_1 + 4x_2 - 6x_3 = 0$$

$$4x_1 + 8x_2 - 10x_3 = 4$$

Corresponding augmented matrix:

$$\left[\begin{array}{ccc|c} 2 & 4 & -6 & 0 \\ 4 & 8 & -10 & 4 \end{array}\right]$$

Corresponding reduced echelon form:

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 6 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

Solution set:

Geometric representation:

EXAMPLE: Compare the solution sets of:

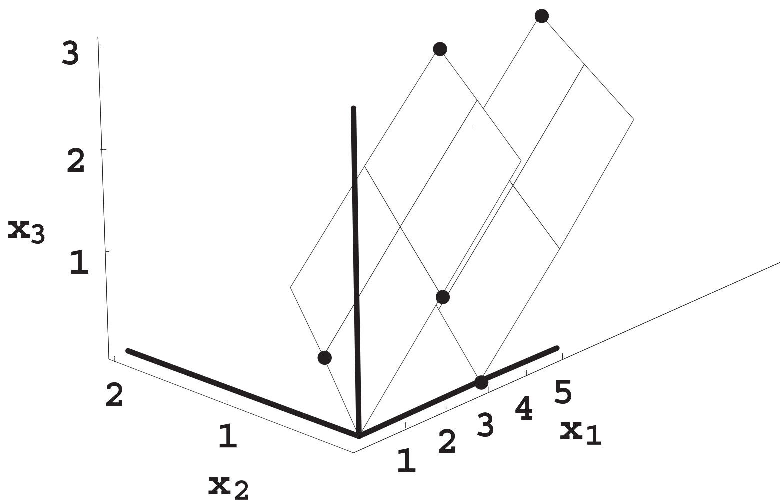
$x_1 - 2x_2 - 2x_3 = 0$ $\qquad\qquad$ $x_1 - 2x_2 - 2x_3 = 3$

Corresponding augmented matrices:

$\left[\begin{array}{ccc|c} 1 & -2 & -2 & 0 \end{array}\right]$ $\qquad\qquad$ $\left[\begin{array}{ccc|c} 1 & -2 & -2 & 3 \end{array}\right]$

These are already in reduced echelon form.
Solution sets:

Geometric representation:



Parallel Solution Sets of $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{x} = \mathbf{b}$

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 $\mathbf{v}$.

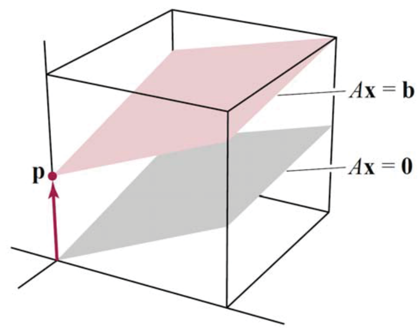
In both cases: to get the solution set of $A\mathbf{x} = \mathbf{b}$, start with the solution set of $A\mathbf{x} = \mathbf{0}$ and translate it by $\mathbf{p}$.

$\mathbf{p}$ 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 Q2b on the exercise sheet:

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 Q2a, 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 Q2b, 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:

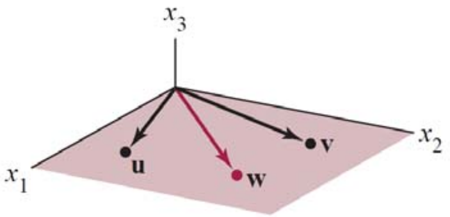
rref $\left(\left[\begin{array}{cccc|c} 1 & 3 & 0 & -4 & 3 \\ 2 & 6 & 0 & -8 & 6 \end{array}\right] \right) = \left[\begin{array}{cccc|c} 1 & 3 & 0 & -4 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$, which gives a different particular solution $\begin{bmatrix} 3 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

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.

§1.7: Linear Independence



In this picture, the plane is Span $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ = Span $\{\mathbf{u}, \mathbf{v}\}$, so we do not need to include $\mathbf{w}$ to describe this plane.
We can think that $\mathbf{w}$ is "too similar" to $\mathbf{u}$ and $\mathbf{v}$ - and linear dependence is the way to make this idea precise.

Definition: A set of vectors $\{\mathbf{v_1}, \dots, \mathbf{v_p}\}$ is *linearly independent* if the only solution to the vector equation

$$x_1\mathbf{v_1} + \dots + x_p\mathbf{v_p} = \mathbf{0}$$

is the trivial solution ($x_1 = \dots = x_p = 0$).

The opposite of linearly independent is linearly dependent:

Definition: A set of vectors $\{\mathbf{v_1}, \dots, \mathbf{v_p}\}$ is *linearly dependent* if there are weights $c_1, \dots, c_p$, not all zero, such that

$$c_1\mathbf{v_1} + \dots + c_p\mathbf{v_p} = \mathbf{0}.$$

The equation $c_1\mathbf{v_1} + \dots + c_p\mathbf{v_p} = \mathbf{0}$ is a linear dependence relation.

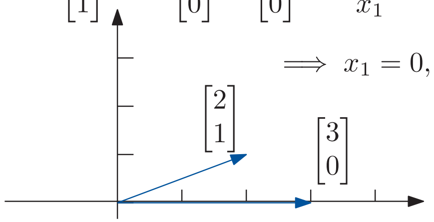
$$x_1\mathbf{v_1} + \dots + x_p\mathbf{v_p} = \mathbf{0}$$

The only solution is $x_1 = \dots = x_p = 0$
$\rightarrow$ linearly independent

Example: $\left\{\begin{bmatrix}2\\1\end{bmatrix}, \begin{bmatrix}3\\0\end{bmatrix}\right\}$ is linearly independent because

$$x_1\begin{bmatrix}2\\1\end{bmatrix} + x_2\begin{bmatrix}3\\0\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix} \Longrightarrow \begin{aligned}2x_1 + 3x_2 &= 0\\ x_1 &= 0\end{aligned}$$

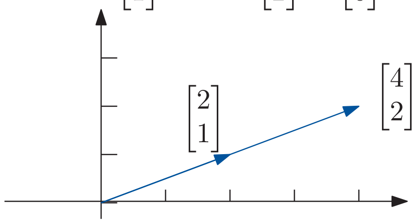
$$\Longrightarrow x_1 = 0, x_2 = 0.$$



There is a solution with some $x_i \neq 0$
$\rightarrow$ linearly dependent

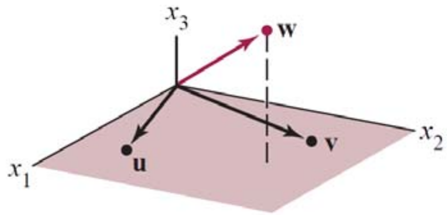
Example: $\left\{\begin{bmatrix}2\\1\end{bmatrix}, \begin{bmatrix}4\\2\end{bmatrix}\right\}$ is linearly dependent because

$$2\begin{bmatrix}2\\1\end{bmatrix} + (-1)\begin{bmatrix}4\\2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}.$$



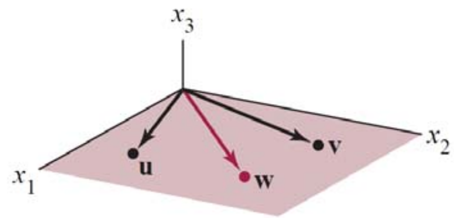
$$x_1\mathbf{v_1} + \dots + x_p\mathbf{v_p} = \mathbf{0}$$

The only solution is $x_1 = \dots = x_p = 0$
(i.e. unique solution)
$\rightarrow$ linearly independent



Informally: $\mathbf{v}_1, \dots, \mathbf{v}_p$ are in "totally different directions"; there is "no relationship" between $\mathbf{v}_1, \dots \mathbf{v}_p$.

There is a solution with some $x_i \neq 0$
(i.e. infinitely many solutions)
$\rightarrow$ linearly dependent



Informally: $\mathbf{v}_1, \dots, \mathbf{v}_p$ are in "similar directions"

Some easy cases:

- Sets containing the zero vector $\{\mathbf{0}, \mathbf{v_2}, \ldots, \mathbf{v_p}\}$:

  $(1)\mathbf{0} + (0)\mathbf{v_2} + \cdots + (0)\mathbf{v_p} = \mathbf{0}$ linearly dependent

- Sets containing one vector $\{\mathbf{v}\}$:

  $x\mathbf{v} = \mathbf{0}$ linearly independent if $\mathbf{v} \neq \mathbf{0}$

$\begin{bmatrix} xv_1 \\ \vdots \\ xv_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$. If some $v_i \neq 0$, then $x = 0$ is the only solution.

Some easy cases:

- Sets containing two vectors $\{\mathbf{u}, \mathbf{v}\}$:

  $x_1\mathbf{u} + x_2\mathbf{v} = \mathbf{0}$

  if $x_1 \neq 0$, then $\mathbf{u} = (-x_2/x_1)\mathbf{v}$. if $x_2 \neq 0$, then $\mathbf{v} = (-x_1/x_2)\mathbf{u}$.

  So $\{\mathbf{u}, \mathbf{v}\}$ is linearly dependent if and only if one of the vectors is a multiple of the other (see p34).

- Sets containing more vectors: $x_1\mathbf{v_1} + \cdots + x_p\mathbf{v_p} = \mathbf{0}$

  A set of vectors is linearly dependent if and only if one of the vectors is a linear combination of the others. (If the weight x_i in the linear dependency relation is non-zero, then $\mathbf{v}_i$ is a linear combination of the other $\mathbf{v}$s.)

How to determine if $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is linearly independent:

EXAMPLE Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 5 \\ 9 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -3 \\ 9 \\ 3 \end{bmatrix}$.

a. Determine if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.
b. If possible, find a linear dependence relation among $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

Solution: (a)

$$x_1\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + x_2\begin{bmatrix} 2 \\ 5 \\ 9 \end{bmatrix} + x_3\begin{bmatrix} -3 \\ 9 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Augmented matrix:

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & 0 \\ 3 & 5 & 9 & 0 \\ 5 & 9 & 3 & 0 \end{array}\right] \quad \text{row reduces to} \quad \left[\begin{array}{ccc|c} 1 & 2 & -3 & 0 \\ 0 & 1 & -18 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

x_3 is a free variable $\Rightarrow$ there are nontrivial solutions.

$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly dependent set

(b) Reduced echelon form: $\left[\begin{array}{ccc|c} 1 & 0 & 33 & 0 \\ 0 & 1 & -18 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

Let $x_3 =$ _____ (any nonzero number). Then $x_1 =$ _____ and $x_2 =$ _____.

$$\_\_\_\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + \_\_\_\begin{bmatrix} 2 \\ 5 \\ 9 \end{bmatrix} + \_\_\_\begin{bmatrix} -3 \\ 9 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$\_\_\_\mathbf{v}_1 + \_\_\_\mathbf{v}_2 + \_\_\_\mathbf{v}_3 = \mathbf{0}$

(one possible linear dependence relation)

A non-trivial solution to $A\mathbf{x} = \mathbf{0}$ is a linear dependence relation between the columns of A: $A\mathbf{x} = \mathbf{0}$ means $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$.

Theorem: Uniqueness of solutions for linear systems: For a matrix A, the following are equivalent:

a. $A\mathbf{x} = \mathbf{0}$ has no non-trivial solution.
b. If $A\mathbf{x} = \mathbf{b}$ is consistent, then it has a unique solution.
c. The columns of A are linearly independent.
d. rref(A) has a pivot in every column (i.e. all variables are basic).

In particular: the row reduction algorithm produces at most one pivot in each row of rref(A). So, if A has more columns than rows (a "fat" matrix), then rref(A) cannot have a pivot in every column.

So a set of more than n vectors in $\mathbb{R}^n$ is always linearly dependent.

Exercise: Combine this with the Theorem of Existence of Solutions (p19) to show that a set of n linearly independent vectors span $\mathbb{R}^n$.

Conceptual problems regarding linear independence:

In problems about linear independence (or spanning) that do not involve specific numbers, it's often better not to compute, i.e. not to use row-reduction.

Example: Prove that, if $\{2\mathbf{u}, \mathbf{v}+\mathbf{w}\}$ is linearly dependent, then $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent.

Method:

Step 1 Rewrite the mathematical terms in the question as formulas. Be careful to distinguish what we know (first line of the proof) and what we want to show (last line of the proof).

What we know: there are scalars c_1, c_2 not both zero such that $c_1(2\mathbf{u}) + c_2(\mathbf{v}+\mathbf{w}) = \mathbf{0}$.

What we want to show: there are scalars d_1, d_2, d_3 not all zero such that $d_1\mathbf{u} + d_2\mathbf{v} + d_3\mathbf{w} = \mathbf{0}$.

(Be careful to choose different letters for the weights in the different statements, because the weights in different statements will in general be different.)

Example: Prove that, if $\{2\mathbf{u}, \mathbf{v}+\mathbf{w}\}$ is linearly dependent, then $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent.

Method:

Step 1 Rewrite the mathematical terms in the question as formulas.

What we know: there are scalars c_1, c_2 not both zero such that $c_1(2\mathbf{u}) + c_2(\mathbf{v}+\mathbf{w}) = \mathbf{0}$.

What we want to show: there are scalars d_1, d_2, d_3 not all zero such that $d_1\mathbf{u} + d_2\mathbf{v} + d_3\mathbf{w} = \mathbf{0}$.

Step 2 Fill in the missing steps by rearranging (and sometimes combining) vector equations.

Answer: We know there are scalars c_1, c_2 not both zero such that

$$c_1(2\mathbf{u}) + c_2(\mathbf{v}+\mathbf{w}) = \mathbf{0}$$
$$2c_1\mathbf{u} + c_2\mathbf{v} + c_2\mathbf{w} = \mathbf{0}$$

and $2c_1, c_2, c_2$ are not all zero, so this is a linear dependence relation among $\mathbf{u}, \mathbf{v}, \mathbf{w}$.

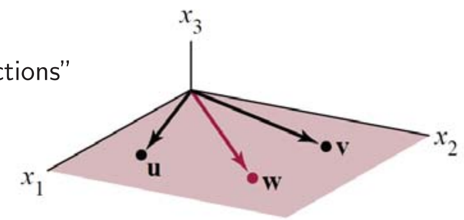
Partial summary of linear dependence:

The definition: $x_1\mathbf{v_1} + \cdots + x_p\mathbf{v_p} = \mathbf{0}$ has a non-trivial solution (not all x_i are zero); equivalently, it has infinitely many solutions.

Equivalently: one of the vectors is a linear combination of the others (see p33, also Theorem 7 in textbook). But it might not be the case that every vector in the set is a linear combination of the others (see Q2c on the exercise sheet).

Computation: rref $\left(\begin{bmatrix} | & & | \\ \mathbf{v}_1 & \dots & \mathbf{v}_p \\ | & & | \end{bmatrix}\right)$ has at least one free variable.

Informal idea: the vectors are in "similar directions"



Partial summary of linear dependence (continued):

Easy examples:

- Sets containing the zero vector;
- Sets containing "too many" vectors (more than n vectors in $\mathbb{R}^n$);
- Multiples of vectors: e.g. $\left\{\begin{bmatrix}2\\1\end{bmatrix}, \begin{bmatrix}4\\2\end{bmatrix}\right\}$ (this is the only possibility if the set has two vectors);
- Other examples: e.g. $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$

Adding vectors to a linearly dependent set still makes a linearly dependent set (see Q2d on exercise sheet).

Equivalent: removing vectors from a linearly independent set still makes a linearly independent set (because P implies Q mean (not Q) implies (not P) - this is the contrapositive).