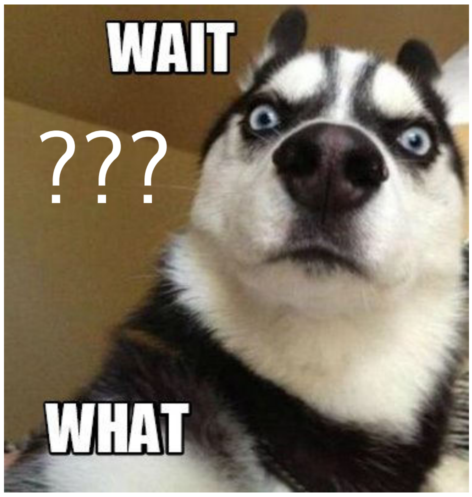
Linear Algebra

Semester 2 2020

1.	§1.1-1.2	linear systems and row reduction
2.	§1.3-1.5	linear combinations, existence of solutions
3.	§1.7	linear independence
4.	§1.8-1.9	linear transformations
5.	§2.1-2.3	matrix operations and inverse
6.	§3.1-3.3	determinants
7.	§4.1-4.2	abstract vector spaces
8.	§4.3-4.4, 5.4	basis and coordinates
.5.	§4.5-4.6	dimension and rank
9.	§4.4, 4.7, 5.4	change of basis
10.	§5.1-5.3	eigenvectors, eigenvalues and diagonalisation
11.	§6.1, 6.5-6.6	length, orthogonality, least squares, regression
12.	§6.2-6.4	orthogonal sets, orthogonal projections, orthogonal matrices
13.	§7.1	diagonalisation of symmetric matrices

Much of the material here comes from previous LinAl teachers at HKBU, and from our official textbook, "Linear Algebra and its Applications" by Lay. Thank you also to my many teachers, co-teachers and students; your views and ideas are incorporated in this retelling.

What is Linear Algebra?



Linear algebra is the study of “adding things”.

In mathematics, there are many situations where we need to “add things” (e.g. numbers, functions, shapes), and linear algebra is about the properties that are common to all these different “additions”. This means we only need to study these properties once, not separately for each type of “addition” (better explanation in Week 7).

Because so many problems require “adding things”, linear algebra is one of the best tools in mathematics.

(picture from mememaker.net)

The concepts in linear algebra are important for many branches of mathematics:

All these classes list Linear Algebra as a prerequisite
(Info from math department website)

Major Requirements for Graduation:

Core Courses (3 units each):

MATH1005 Calculus I

MATH2225 Calculus II

MATH2205 Multivariate Calculus

MATH2206 Probability & Statistics

MATH2207 Linear Algebra

MATH2215 Mathematical Analysis

MATH2216 Statistical Methods and Theory

MATH3205 Linear Programming and Integer Programming

MATH3206 Numerical Methods I

MATH3405 Ordinary Differential Equations

MATH3805 Regression Analysis

MATH3806 Multivariate Statistical Methods

MATH4998 Mathematical Science Project I

This class is about more than calculations. From the official syllabus:

Course Intended Learning Outcomes (CILOs):

Upon successful completion of this course, students should be able to:

No.	**Course Intended Learning Outcomes (CILOs)**
1	Explain the concept/theory in linear algebra, to develop dynamic and graphical views to the related issues of the chosen topics as outlined in "course content," and to formally prove theorems

Linear algebra is used in future courses in entirely different ways. So it's not enough to know routine calculations; you need to understand the concepts and ideas, to solve problems you haven't seen before on the exam. This will require words and not just formulae.

For many people, this is different from their previous math classes, and will require a lot of study time.
(Week 1 is straightforward computation; the abstract theory starts in Week 2.)

§1.1: Systems of Linear Equations

Linear Algebra starts with linear equations.

Example: $y = 5x + 2$ is a linear equation. We can take all the variables to the left hand side and rewrite this as $(-5)x + (1)y = 2$.

Example: $3(x_1 + 2x_2) + 1 = x_1 + 1 \longrightarrow (2)x_1 + (6)x_2 = 0$

Example: $x_2 = \sqrt{2}(\sqrt{6} - x_1) + x_3 \longrightarrow \sqrt{2}x_1 + (1)x_2 + (-1)x_3 = 2\sqrt{3}$

The following two equations are not linear, why?

$x_2 = 2\sqrt{x_1}$ $\qquad$ $xy + x = e^5$ $\qquad$ The problem is that the variables are not only multiplied by numbers.

In general, a linear equation is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

$x_1, x_2, \ldots x_n$ are the variables. $\qquad$ $a_1, a_2, \ldots a_n$ are the coefficients.

A linear equation has the form $a_1x_1 + a_2x_2 + \ldots a_nx_n = b$.

Definition: A *system of linear equations* (or a *linear system*) is a collection of linear equations involving the same set of variables.

Example:
$$\begin{array}{rrrcr} x & +y & & = & 3 \\ 3x & & +2z & = & -2 \end{array}$$
is a system of 2 equations in 3 variables, x, y, z. Notice that not every variable appears in every equation.

Definition: A *solution* of a linear system is a list $(s_1, s_2, \ldots, s_n)$ of numbers that makes each equation a true statement when the values $s_1, s_2, \ldots, s_n$ are substituted for $x_1, x_2, \ldots, x_n$ respectively.

Definition: The *solution set* of a linear system is the set of all possible solutions.

Example: One solution to the above system is $(x, y, z) = (2, 1, -4)$, because $2 + 1 = 3$ and $3(2) + 2(-4) = -2$.

Question: Is there another solution? How many solutions are there?

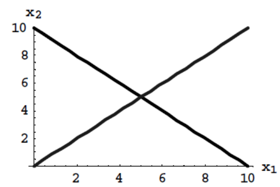
Definition: A linear system is *consistent* if it has a solution,
and *inconsistent* if it does not have a solution.

Fact: (which we will prove in the next class) A linear system has either

- exactly one solution — consistent
- infinitely many solutions — consistent
- no solutions — inconsistent

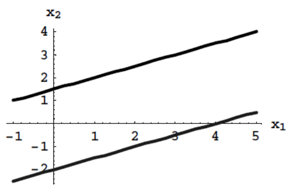
EXAMPLE Two equations in two variables:

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



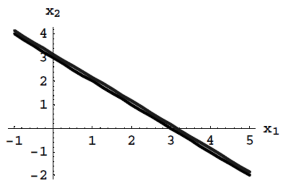
one unique solution
consistent

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



no solution
inconsistent

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

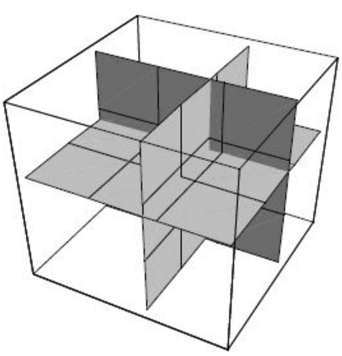


infinitely many solutions
consistent

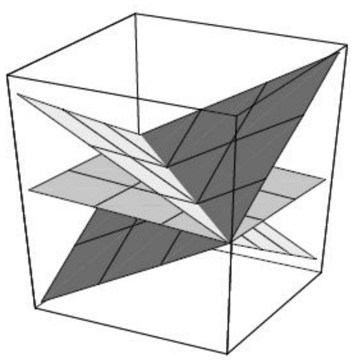
i.e. $ax + by + cz = d$

EXAMPLE: Three equations in three variables. Each equation determines a plane in 3-space.

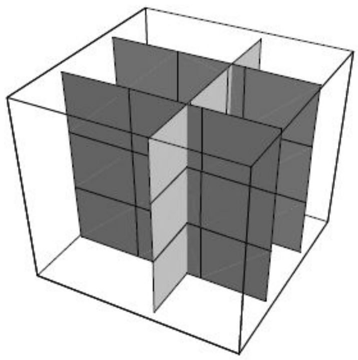
i) The planes intersect in one point. *(one solution)*



ii) The planes intersect in one line. *(infinitely many solutions)*



iii) There is no point in common to all three planes. *(no solution)*

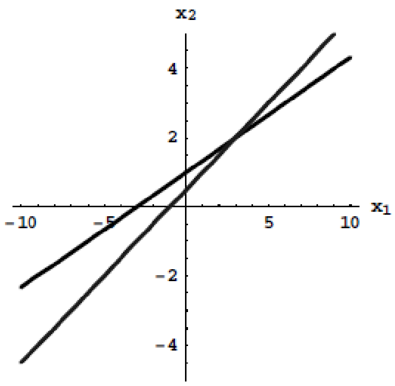


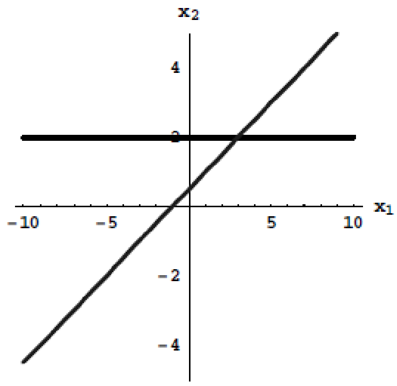
Which of these cases are consistent?

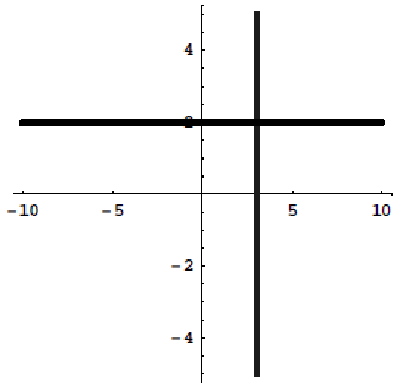
consistent consistent inconsistent

Our goal for this week is to develop an efficient algorithm to solve a linear system.
Example:

$$\begin{array}{lrcrcr} R_1 & x_1 & - & 2x_2 & = & -1 \\ R_2 & -x_1 & + & 3x_2 & = & 3 \end{array} \quad \begin{array}{r} \rightarrow \\ R_2 + R_1 \rightarrow \end{array} \quad \begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ & & x_2 & = & 2 \end{array} \quad \begin{array}{r} R_1 + 2R_2 \rightarrow \\ \end{array} \quad \begin{array}{rcr} x_1 & = & 3 \\ x_2 & = & 2 \end{array}$$







Definition: Two linear systems are *equivalent* if they have the same solution set.

So the three linear systems above are different but equivalent.
A general strategy for solving a linear system: replace one system with an equivalent system that is easier to solve.

We simplify the writing by using matrix notation, recording only the coefficients and not the variables.

$$\begin{array}{lrcrcr} R_1 & x_1 & - & 2x_2 & = & -1 \\ R_2 & -x_1 & + & 3x_2 & = & 3 \end{array} \quad \begin{array}{r} \rightarrow \\ R_2 + R_1 \rightarrow \end{array} \quad \begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ & & x_2 & = & 2 \end{array} \quad \begin{array}{r} R_1 + 2R_2 \rightarrow \\ \end{array} \quad \begin{array}{rcr} x_1 & = & 3 \\ x_2 & = & 2 \end{array}$$

$$\left[\begin{array}{rr|r} 1 & -2 & -1 \\ -1 & 3 & 3 \end{array}\right] \longrightarrow \left[\begin{array}{rr|r} 1 & -2 & -1 \\ 0 & 1 & 2 \end{array}\right] \longrightarrow \left[\begin{array}{rr|r} 1 & 0 & 3 \\ 0 & 1 & 2 \end{array}\right]$$

coefficient of x_1 coefficient of x_2 right hand side

The augmented matrix of a linear system contains the right hand side: $\left[\begin{array}{rr|r} 1 & -2 & -1 \\ -1 & 3 & 3 \end{array}\right]$

The coefficient matrix of a linear system is the left hand side only: $\left[\begin{array}{rr} 1 & -2 \\ -1 & 3 \end{array}\right]$

(The textbook does not put a vertical line between the coefficient matrix and the right hand side, but I strongly recommend that you do to avoid confusion.)

$$\begin{array}{lrcrcr} R_1 & x_1 & - & 2x_2 & = & -1 \\ R_2 & -x_1 & + & 3x_2 & = & 3 \end{array} \quad \xrightarrow{R_2+R_1\to} \quad \begin{array}{rcrcr} x_1 & - & 2x_2 & = & -1 \\ & & x_2 & = & 2 \end{array} \quad \xrightarrow{R_1+2R_2\to} \quad \begin{array}{rcr} x_1 & = & 3 \\ x_2 & = & 2 \end{array}$$

$$\left[\begin{array}{rr|r} 1 & -2 & -1 \\ -1 & 3 & 3 \end{array}\right] \longrightarrow \left[\begin{array}{rr|r} 1 & -2 & -1 \\ 0 & 1 & 2 \end{array}\right] \longrightarrow \left[\begin{array}{rr|r} 1 & 0 & 3 \\ 0 & 1 & 2 \end{array}\right]$$

In this example, we solved the linear system by applying elementary row operations to the augmented matrix (we only used 1. above, the others will be useful later):

1. Replacement: add a multiple of one row to another row. $R_i \to R_i + cR_j$
2. Interchange: interchange two rows. $R_i \to R_j,\ R_j \to R_i$
3. Scaling: multiply all entries in a row by a nonzero constant. $R_i \to cR_i,\ c \neq 0$

Definition: Two matrices are *row equivalent* if one can be transformed into the other by a sequence of elementary row operations.

Fact: If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set, i.e. they are equivalent linear systems.

General strategy for solving a linear system: do row operations to its augmented matrix to get an equivalent system that is easier to solve.

EXAMPLE:

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & 2x_2 & - & 8x_3 & = & 8 \\ -4x_1 & + & 5x_2 & + & 9x_3 & = & -9 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array}\right]$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & 2x_2 & - & 8x_3 & = & 8 \\ & - & 3x_2 & + & 13x_3 & = & -9 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ & & & \end{array}\right]$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & x_2 & - & 4x_3 & = & 4 \\ & - & 3x_2 & + & 13x_3 & = & -9 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{array}\right]$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & x_2 & - & 4x_3 & = & 4 \\ & & & & x_3 & = & \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ & & & \end{array}\right]$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & & & = & -3 \\ & & x_2 & & & = & \\ & & & & x_3 & = & 3 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -2 & 0 & -3 \\ & & & \\ 0 & 0 & 1 & 3 \end{array}\right]$$

$$\begin{array}{rcrcrcr} x_1 & & & & & = & \\ & & x_2 & & & = & 16 \\ & & & & x_3 & = & 3 \end{array} \qquad \left[\begin{array}{rrr|r} & & & \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Solution: $(x_1,x_2,x_3)=(29,16,3)$

Check: Is $(29,16,3)$ a solution of the ***original*** system?

Warning: Do not do multiple elementary row operations at the same time, except adding multiples of the same row to several rows.

$$\begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 3x_2 &= 3 \end{aligned} \qquad \begin{aligned} x_2 &= 2 \\ x_2 &= 2 \end{aligned}$$

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

These are NOT equivalent systems: in the system on the right, x_1 can take any value, which is not true for the system on the left.

$$\begin{aligned} x_1 - 2x_2 &= -3 \\ x_2 &= 16 \\ x_3 &= 3 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array}\right] \begin{array}{l} \leftarrow R_1 - R_3 \\ \leftarrow R_2 + 4R_3 \\ \end{array}$$

$$\begin{aligned} x_1 &= 29 \\ x_2 &= 16 \\ x_3 &= 3 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Sometimes we are not interested in the exact value of the solutions, just the number of solutions. In other words:

1. Existence of solutions: is the system consistent?
2. Uniqueness of solutions: if a solution exists, is it the only one?

Answering this requires less work than finding the solution.

Example:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array}\right]$$

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{array}\right]$$

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{array}\right]$$

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ x_3 &= 3 \end{aligned} \qquad \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

We can stop here: back-substitution shows that we can find a unique solution.

EXAMPLE: Is this system consistent?

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= -1 \\ 5x_1 - 7x_2 + 9x_3 &= 0 \\ 3x_2 - 6x_3 &= 8 \end{aligned}$$

EXAMPLE: For what values of h will the following system be consistent?

$$x_1 - 3x_2 = 4$$

$$-2x_1 + 6x_2 = h$$

Section 1.2: Row Reduction and Echelon Forms

Motivation: it is easy to solve a linear system whose augmented matrix is in reduced echelon form

Echelon form (or **row echelon form):**

1. All nonzero rows are above any rows of all zeros.
2. Each *leading entry* (i.e. left most nonzero entry) of a row is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zero.

EXAMPLE: Echelon forms

(a) $\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & \blacksquare \\ 0 & 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \end{bmatrix}$

Reduced echelon form: Add the following conditions to conditions 1, 2, and 3 above:

4. The leading entry in each nonzero row is 1.
5. Each leading 1 is the only nonzero entry in its column.

EXAMPLE (continued):

Reduced echelon form :

$$\begin{bmatrix} 0 & 1 & * & 0 & 0 & * & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 1 & 0 & * & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \end{bmatrix}$$

EXAMPLE: Are these matrices in echelon form, reduced echelon form, or neither?

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & 2x_2 & - & 8x_3 & = & 8 \\ -4x_1 & + & 5x_2 & + & 9x_3 & = & -9 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array}\right]$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & 2x_2 & - & 8x_3 & = & 8 \\ & - & 3x_2 & + & 13x_3 & = & -9 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{array}\right]$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & x_2 & - & 4x_3 & = & 4 \\ & - & 3x_2 & + & 13x_3 & = & -9 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{array}\right]$$

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & x_3 & = & 0 \\ & & x_2 & - & 4x_3 & = & 4 \\ & & & & x_3 & = & 3 \end{array} \qquad \boxed{\left[\begin{array}{rrr|r} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right]}$$

echelon form

$$\begin{array}{rcrcr} x_1 & - & 2x_2 & = & -3 \\ & & x_2 & = & 16 \\ & & x_3 & = & 3 \end{array} \qquad \left[\begin{array}{rrr|r} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

$$\begin{array}{rcr} x_1 & = & 29 \\ x_2 & = & 16 \\ x_3 & = & 3 \end{array} \qquad \boxed{\left[\begin{array}{rrr|r} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array}\right]}$$

reduced echelon form

Here is the example from p10. Notice that we use row operations to first put the matrix into echelon form, and then into reduced echelon form.

Can we always do this for any linear system?

Theorem: Any matrix A is row-equivalent to exactly one reduced echelon matrix, which is called its reduced echelon form and written rref(A).

So our general strategy for solving a linear system is: apply row operations to its augmented matrix to obtain its rref.

And our general strategy for determining existence/uniqueness of solutions is: apply row operations to its augmented matrix to obtain an echelon form, i.e. a row-equivalent echelon matrix.

Warning: an echelon form is not unique. Its entries depend on the row operations we used. But its pattern of $\blacksquare$ and $*$ is unique.

These processes of row operations (to get to echelon or reduced echelon form) are called row reduction.

Row reduction:

augmented matrix of linear system $\longrightarrow$ echelon form $\longrightarrow$ reduced echelon form

echelon form $\downarrow$ existence and uniqueness of solutions

reduced echelon form $\downarrow$ solution set

The rest of this section:
- The row reduction algorithm (p21-25);
- Getting the solution, existence/uniqueness from the (reduced) echelon form (p26-29).

Important terms in the row reduction algorithm:
- pivot position: the position of a leading entry in a row-equivalent echelon matrix.
- pivot: a nonzero entry of the matrix that is used in a pivot position to create zeroes below it.
- pivot column: a column containing a pivot position.

The black squares are the pivot positions.

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \end{bmatrix}$$

Row reduction algorithm:

EXAMPLE:

$$\left[\begin{array}{ccccc|c} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 1 & -3 & 4 & -3 & 2 & 5 \end{array}\right]$$

1. The top of the leftmost nonzero column is a pivot position.

2. Put a pivot in this position, by scaling or interchanging rows.

$$\begin{array}{cc} \left[\begin{array}{ccccc|c} 1 & -3 & 4 & -3 & 2 & 5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right] & \begin{array}{c} R_3 \\ \\ R_1 \end{array} \end{array}$$

3. Create zeroes in all positions below the pivot, by adding multiples of the top row to each row.

$$\left[\begin{array}{ccccc|c} 1 & -3 & 4 & -3 & 2 & 5 \\ & & & & & \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right]$$

4. Ignore this row and all rows above, and repeat steps 1-3.

$$\left[\begin{array}{ccccc|c} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right]$$

1. The top of the leftmost nonzero column is a pivot position.
2. Put a pivot in this position, by scaling or interchanging rows.

$$\left[\begin{array}{ccccc|c} 1 & -3 & 4 & -3 & 2 & 5 \\ & & & & & \\ 0 & 3 & -6 & 6 & 4 & -5 \end{array}\right]$$

3. Create zeroes in all positions below the pivot, by adding multiples of the top row to each row.

$$\left[\begin{array}{ccccc|c} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ & & & & & \end{array}\right]$$

4. Ignore this row and all rows above, and repeat steps 1-3.

$$\left[\begin{array}{ccccc|c} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right]$$

1. The top of the leftmost nonzero column is a pivot position.
2. Put a pivot in this position, by scaling or interchanging rows.
3. Create zeroes in all positions below the pivot, by adding multiples of the top row to each row.

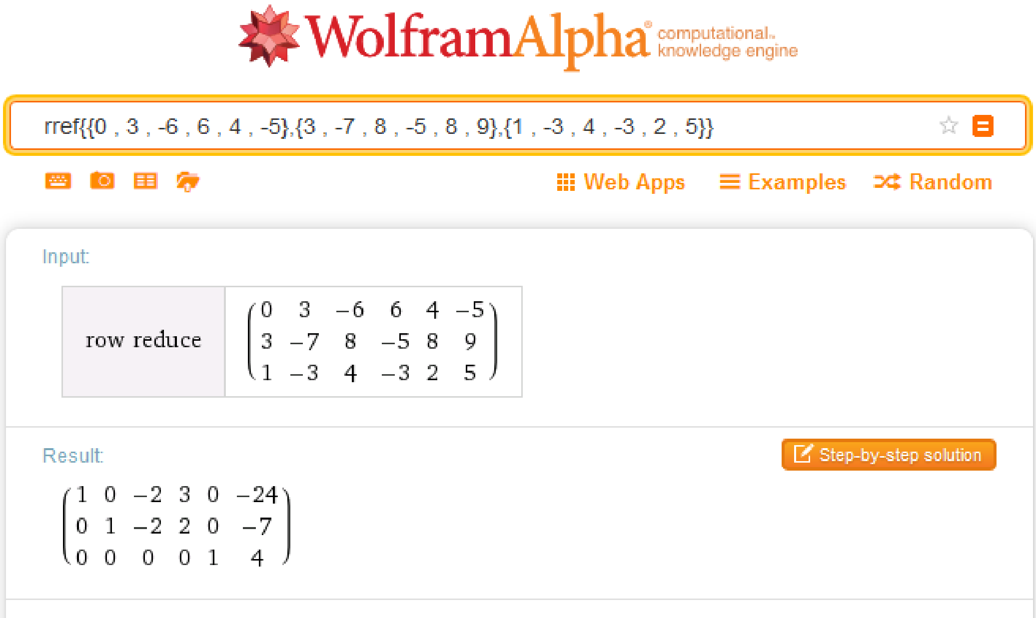
We are at the bottom row, so we don't need to repeat anymore. We have arrived at an echelon form.

5. To get from echelon to reduced echelon form (back substitution):
   Starting from the bottom row: for each pivot, add multiples of the row with the pivot to the other rows to create zeroes above the pivot.

$$\left[\begin{array}{ccccc|c} 1 & -3 & 4 & -3 & 0 & -3 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right] \begin{array}{c} R_1 - 2R_3 \\ R_2 - R_3 \\ \end{array}$$

$$\left[\begin{array}{ccccc|c} & & & & & \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right]$$

Check your answer: www.wolframalpha.com



Getting the solution set from the reduced echelon form:

A basic variable is a variable corresponding to a pivot column.
All other variables are free variables.

6. Write each row of the augmented matrix as a linear equation.

Example:

$$\left[\begin{array}{ccccc|c} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right] \qquad \begin{aligned} x_1 \quad -2x_3 + 3x_4 \quad &= -24 \\ x_2 - 2x_3 + 2x_4 \quad &= -7 \\ x_5 &= 4 \end{aligned}$$

basic variables: x_1, x_2, x_5, free variables: x_3, x_4.

The free variables can take any value. These values then uniquely determine the basic variables.

7. Take the free variables in the equations to the right hand side, and add equations of the form "free variable = itself", so we have equations for each variable in terms of the free variables.

Example:

$$\begin{aligned} x_1 &= -24 + 2x_3 - 3x_4 \\ x_2 &= -7 + 2x_3 - 2x_4 \\ x_3 &= x_3 \\ x_4 &= x_4 \\ x_5 &= 4 \end{aligned}$$

So the solution set is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -24 + 2s - 3t \\ -7 + 2s - 2t \\ s \\ t \\ 4 \end{pmatrix}$$

where s and t can take any value.

What this means: for every choice of s and t, we get a different solution:
e.g. $s = 0, t = 1$: $(x_1, x_2, x_3, x_4, x_5) = (-27, -9, 0, 1, 4)$
$s = 1, t = -1$: $(x_1, x_2, x_3, x_4, x_5) = (-19, -3, 1, -1, 4)$
and infinitely many others. (Exercise: check these two are solutions.)

We will see a better way to write the solution set next week (Week 2 p29-31, §1.5).

Answering existence and uniqueness of solutions from the echelon form

Example: On p14 we found
$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & -1 \\ 5 & -7 & 9 & 0 \\ 0 & 3 & -6 & 8 \end{array}\right] \xrightarrow{\text{row-reduction}} \left[\begin{array}{ccc|c} \boxed{1} & -2 & 3 & -1 \\ 0 & \boxed{3} & -6 & 5 \\ 0 & 0 & 0 & \boxed{3} \end{array}\right]$$

The last equation says $0x_1 + 0x_2 + 0x_3 = 3$, so this system is inconsistent. Generalising this observation gives us "half" of the following theorem:

Theorem 2: Existence and Uniqueness:
A linear system is consistent if and only if an echelon form of its augmented matrix has no row of the form $[0 \dots 0|\blacksquare]$ with $\blacksquare \neq 0$.

Be careful with the logic here: this theorem says "if and only if", which means it claims two different things:

- If a linear system is consistent, then an echelon form of its augmented matrix cannot contain $[0 \dots 0|\blacksquare]$ with $\blacksquare \neq 0$.
  This is the observation from the example above.

- If there is no row $[0 \dots 0|\blacksquare]$ with $\blacksquare \neq 0$ in an echelon form of the augmented matrix, then the system is consistent.
  This is because we can continue the row-reduction to the rref, and then the solution method of p26-27 will give us solutions.

As for the uniqueness of solutions:

Theorem 2: Existence and Uniqueness:
If a linear system is consistent, then:
- it has a unique solution if there are no free variables;
- it has infinitely many solutions if there are free variables.

In particular, this proves the fact we saw earlier, that a linear system has either a unique solution, infinitely many solutions, or no solutions.

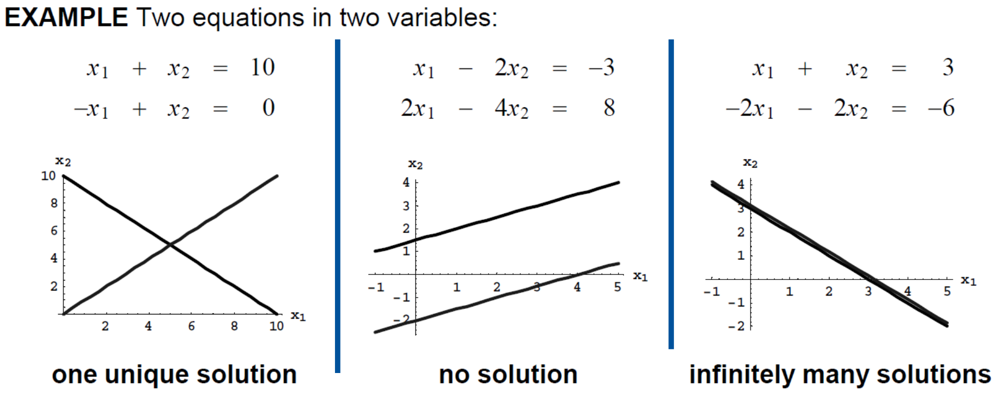
Warning: In general, the existence of solutions is unrelated to the uniqueness of solutions. (We will meet an important exception in §2.3.)

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.



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

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:

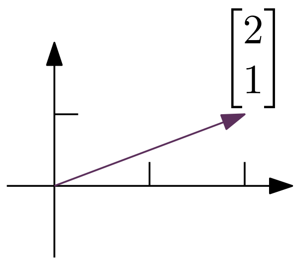
- The related computations: to solve problems about a specific linear system with numbers (Week 2 p10, Week 3 p9-10).
- The rigorous definition: to prove statements about an abstract linear system (Week 2 p15, Week 3 p13).
- The conceptual idea: to guess whether statements are true, to develop a plan for a proof or counterexample, and to help you remember the main theorems (Week 2 p13-14, Week 3 p3-5). This informal view is for thinking only, NOT for answering problems on homeworks and exams.

§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 $c\mathbf{u} = \begin{bmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{bmatrix}$.

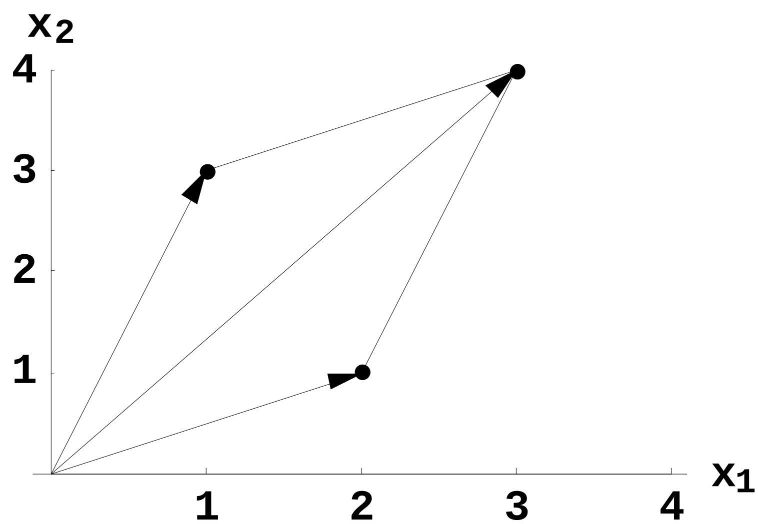
These satisfy the usual rules for arithmetic of numbers, e.g.

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

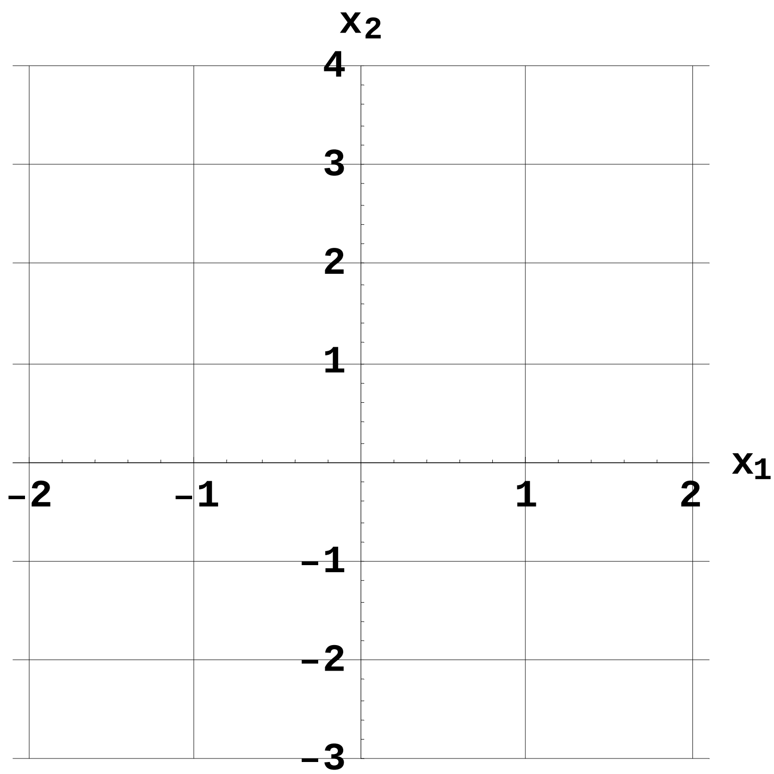
Parallelogram rule for addition of two vectors:

If **u** and **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 **0**, **u** and **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 $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ 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:

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

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

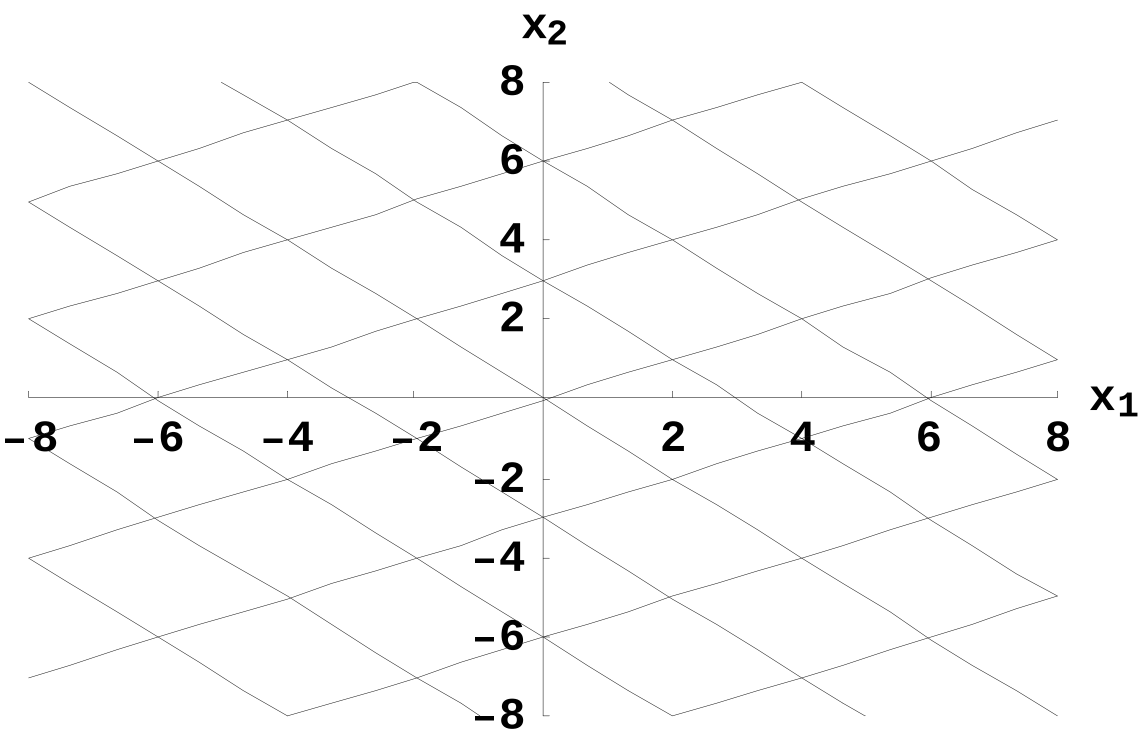
(i.e. $\mathbf{u} + (-3)\mathbf{v}$)

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 $\mathbf{v}_1, \ldots, \mathbf{v}_p$".

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 $\mathbf{v}_2$:

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



When we don't have the grid paper:

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

Express $\mathbf{b}$ as a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$.

Solution: Vector $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$ if ______________________________

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:

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

Exercise: Use this algebraic method on the examples on the previous page and check that you get the same answer.

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

In other words, $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ 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"

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\} = \{\underbrace{x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p}_{\text{vectors of the form } x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p} \mid \underbrace{x_1, \ldots, x_p \in \mathbb{R}}_{\substack{x_1, \ldots x_p \text{ are real numbers} \\ \text{(i.e. they can take any value)}}}\}.$$

the set of

such that

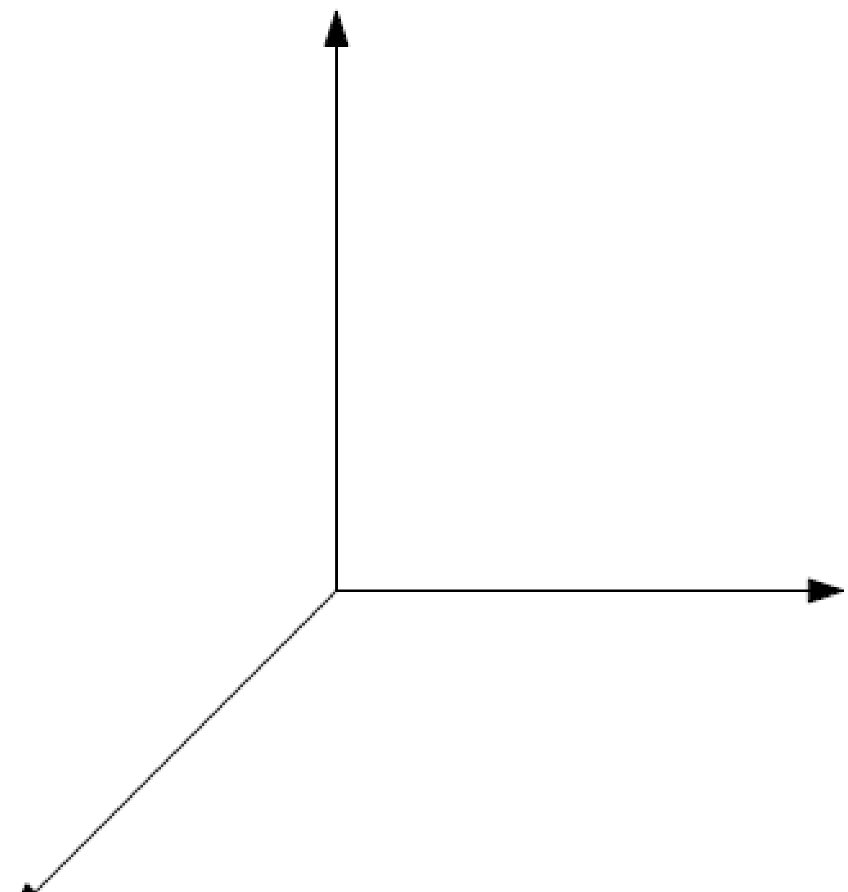
DEFINITION: $\textbf{Span}\,\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\} = \{x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_p\mathbf{v}_p \mid x_1, \ldots, x_p \in \mathbb{R}\}$

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

When $p = 1$, the definition says $\mathrm{Span}\,\{\mathbf{v}_1\} = \{x_1\mathbf{v}_1 \mid x_1 \in \mathbb{R}\}$,

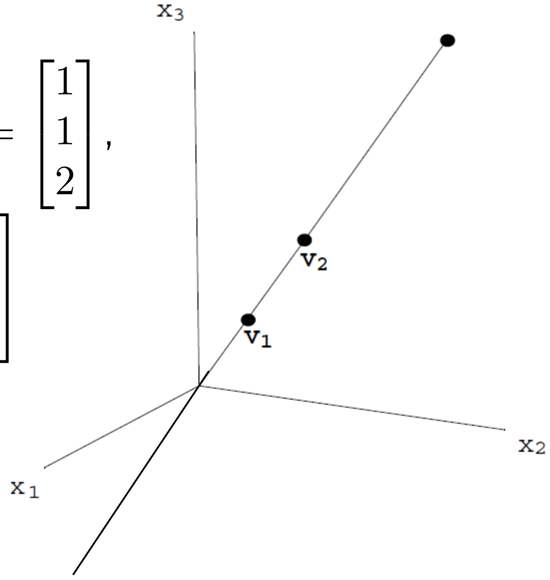
i.e. $\mathrm{Span}\,\{\mathbf{v}_1\}$ is all scalar multiples of $\mathbf{v}_1$.

- $\mathrm{Span}\,\{\mathbf{0}\} = \{\mathbf{0}\}$, because $x_1\mathbf{0} = \mathbf{0}$ for all scalars x_1.
- If $\mathbf{v}_1$ is not the zero vector, then $\mathrm{Span}\,\{\mathbf{v}_1\}$ is ______________________________



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

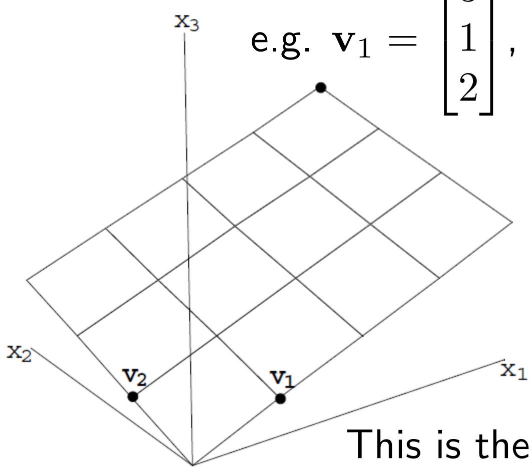


$\mathbf{v}_2$ is a multiple of $\mathbf{v}_1$

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

A first exercise in writing proofs.

Each proof is different. Here are some general guidelines, but not every proof is like this. In particular, do NOT memorise and copy the equations in a particular proof, it will NOT work for a different question.

EXAMPLE: Prove that, if $\mathbf{u}$ is in $\mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, then $2\mathbf{u}$ is in $\mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

STRATEGY:

Step 1: Find the conclusion of the proof, i.e. what the question is asking for. Using definitions (ctrl-F in the notes if you don't remember), write it out as formulas:

Step 2: **on a separate piece of paper**, use definitions to write out the given information as formulas. Be careful to use different letters in different formulas.

Step 3: If the required conclusion (from Step 1) is an equation: start with the left hand side, and calculate/reorganise it using the information in Step 2 to obtain the right hand side.

(More examples: week 3 p13, week 4 p22, week 5 p17, week 5 p19, many exercise sheets.)

The professional way to write this (which may be confusing for beginners):

In more complicated proofs, you may want to use theorems (see week 5 p26).

To improve your proofs:

- **Memorise your definitions**, i.e. how to translate a technical term into a formula.
- After finishing a proof, think about why that strategy works, and why other strategies that you tried didn't work.
- Come to office hours with questions from homework or the textbook, we can do them together.

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.

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

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 $\mathbf{x}$. 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. And do not write $A \cdot \mathbf{x}$, that 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 $A\mathbf{x} = \mathbf{b}$.

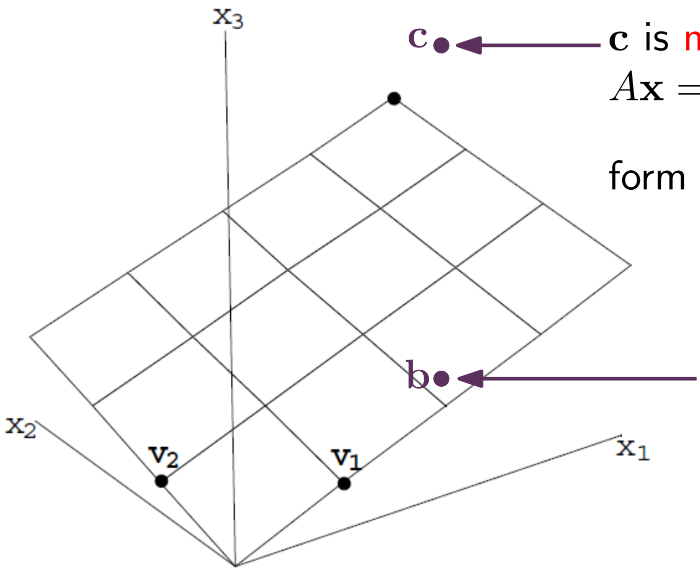
These three problems have the same solution set, so the following three things are the same (they are simply different ways to say "the above problem has a solution"):

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



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

$\mathbf{b}$ 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 p29-31 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 $\mathbb{R}^3$.

Because $\mathbf{b} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2$, the way to see a solution $\mathbf{x}$ on this $\mathbb{R}^3$ picture is like on p9: $\mathbf{x}$ gives the location of $\mathbf{b}$ relative to the gridlines drawn by $\mathbf{v}_1$ and $\mathbf{v}_2$, i.e x_i tells you how far $\mathbf{b}$ is in the $\mathbf{v}_i$ direction (see week 8 p22). For example, for the lower purple dot, $x_1 \sim 2.2$ and $x_2 \sim 0.2$.

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. A has a pivot position in every row.

You may view d) as a computation (reduction to echelon form) to check for a), b) or c).

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. (This is the same as "if (a) is true, then (d) is true".)

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
d. A has a pivot position 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 $[0 \dots 0|\blacksquare]$, 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. A has a pivot position 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.

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

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

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. A has a pivot position in every row.

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

Observe that A has at most one pivot position 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 A cannot have a pivot position 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?

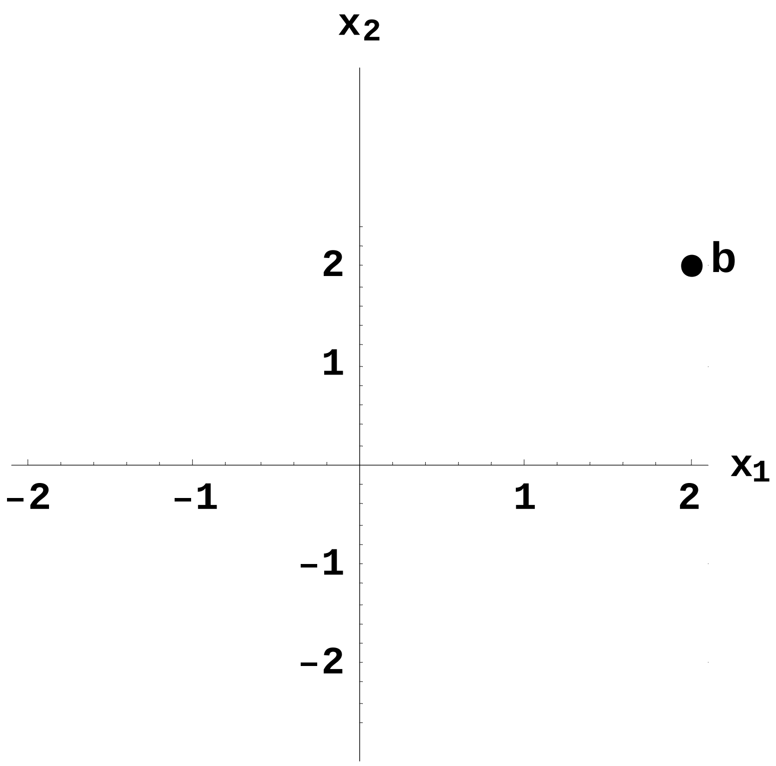
Remember that the solutions to $A\mathbf{x} = \mathbf{b}$ are the weights for writing $\mathbf{b}$ as a linear combination of the columns of A.

$$A\mathbf{x} = \mathbf{b} \quad \longleftrightarrow \quad \mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_p\mathbf{a}_p.$$

The "linear combination of columns" viewpoint gives us a picture way to understand existence of solutions (p20).

Here is a picture about uniqueness of solutions: what does it mean for the weights x_i to be non-unique.

EXAMPLE: $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{a}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$.



Informally, the non-uniqueness of weights happens because we can use our given vectors to "walk in a circle back to $\mathbf{0}$" - this is the idea of linear dependence (week 3).

§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 $[0 \dots 0|\blacksquare]$.
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:

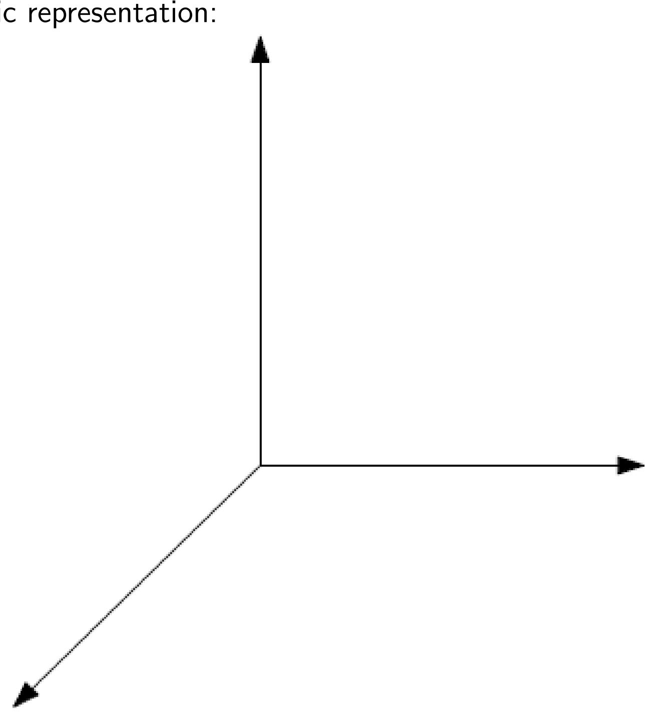
$$\left[\begin{array}{ccc|c} 2 & 4 & -6 & 0 \\ 4 & 8 & -10 & 0 \end{array}\right]$$

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:

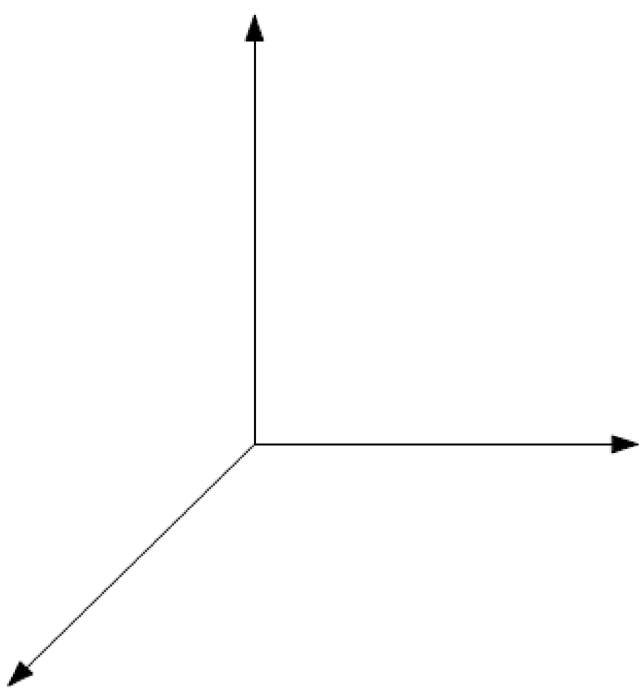
$$\left[\begin{array}{ccc|c} 2 & 4 & -6 & 0 \\ 4 & 8 & -10 & 4 \end{array}\right]$$

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:

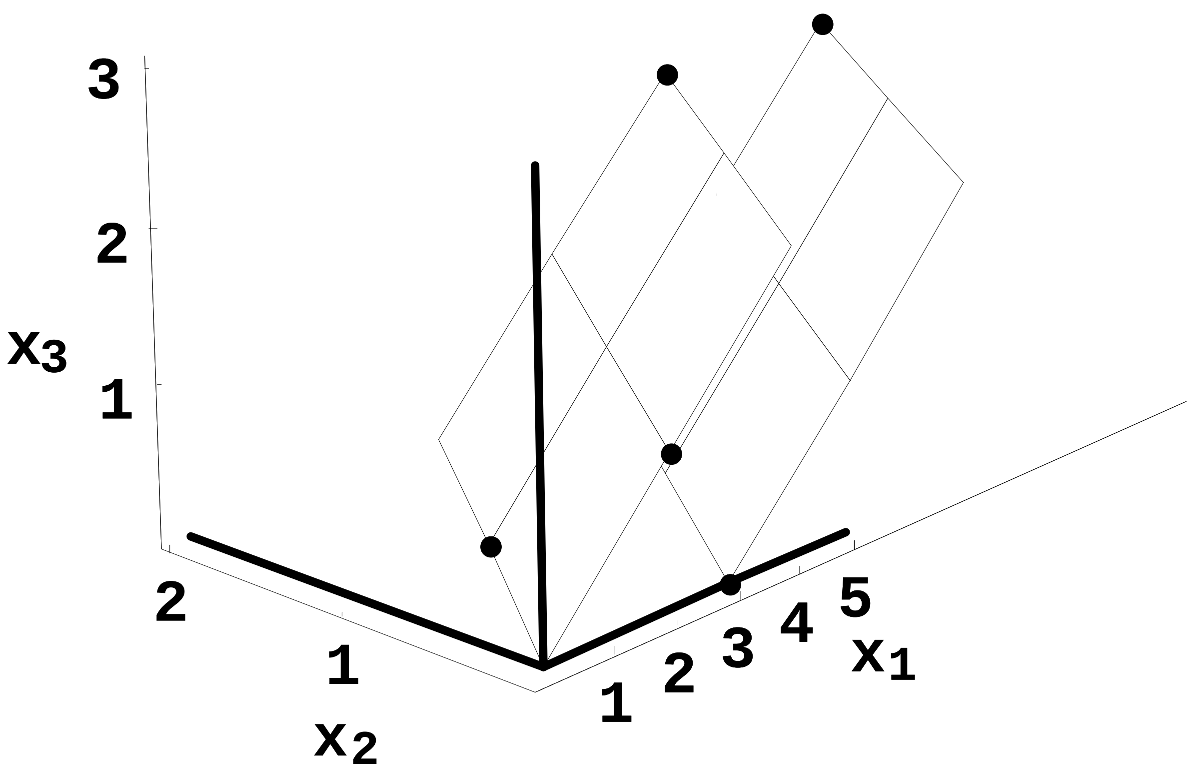
$$x_1 - 2x_2 - 2x_3 = 0 \qquad\qquad x_1 - 2x_2 - 2x_3 = 3$$

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 $\mathbf{v}$.

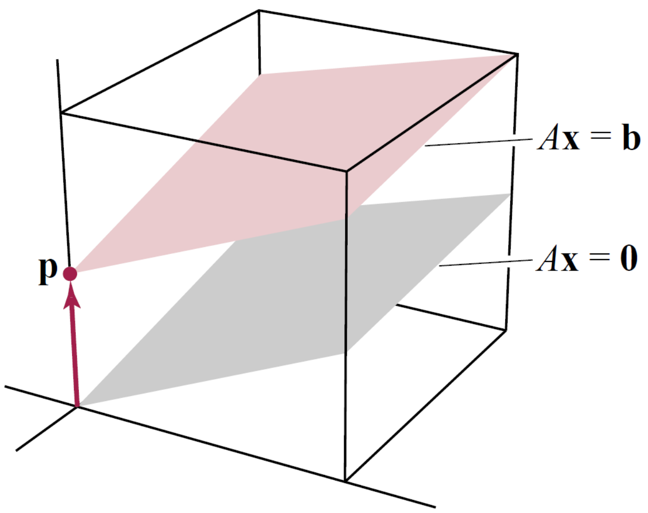
In both cases: to get the solution set of $A\mathbf{x} = \mathbf{b}$, start with the solution set of $A\mathbf{x} = \mathbf{0}$ and translate it by $\mathbf{p}$.

$\mathbf{p}$ 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.

Two typical applications of this theorem:

1. If you write the solutions to $A\mathbf{x} = \mathbf{b}$ in parametric form, then the part with free variables is the solution to $A\mathbf{x} = \mathbf{0}$, e.g. on week1 p26, we found that the

$$\text{solutions to } \begin{bmatrix} 0 & 3 & -6 & 6 & 4 \\ 3 & -7 & 8 & -5 & 8 \\ 1 & -3 & 4 & -3 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} -5 \\ 9 \\ 5 \end{bmatrix} \text{ is } \underbrace{\boxed{\begin{bmatrix} -24 \\ -7 \\ 0 \\ 0 \\ 4 \end{bmatrix}}}_{\text{particular solution}} + \underbrace{\boxed{s \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}}}_{\text{solutions to } \begin{bmatrix} 0 & 3 & -6 & 6 & 4 \\ 3 & -7 & 8 & -5 & 8 \\ 1 & -3 & 4 & -3 & 2 \end{bmatrix} \mathbf{x} = \mathbf{0}}$$

2. If you already have the solutions to $A\mathbf{x} = \mathbf{0}$ and you need to solve $A\mathbf{x} = \mathbf{b}$, then you don't need to row-reduce again: simply find one particular solution (e.g. by guessing) and then add it to the solution set to $A\mathbf{x} = \mathbf{0}$ (example on next page).

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:

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

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

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