Matrix Algebra

Amir Esteghamatian

March 27, 2017

1 Linear systems

A equation is called **linear** when it follows the following form:

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b \tag{1}$$

Example 1.1.

Example 1.2.

Example 1.3.

Example 1.4.

A set of linear equations with the same variables is called a **linear system**:

$$x_1 + 2x_2 - 1.5x_3 = 8 \tag{2}$$

$$2x_1 - x_3 = 0 \tag{3}$$

A solution is a list of numbers which makes each equation a true statement. For above set of linear equations for instance, (3,7,6) is a solution since by substituting it to (2) and (3), it yields $8 = 8$ and $0 = 0$. For this particular example, another solution is (4,8,8), meaning that the solution is not unique. All sets of numbers satisfying a set of linear equations are called the **solution set** for that set of linear equations. Two linear sets are equivalent if they have the same solution set.

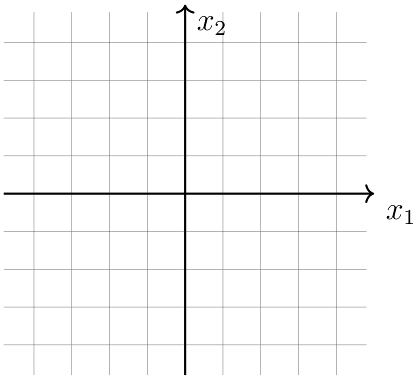
For a better visualization of a set of linear equations, let's start with a simple set of two equations with two variables:

Example 1.5.

$$x_2 + x_1 = 1 \tag{4}$$

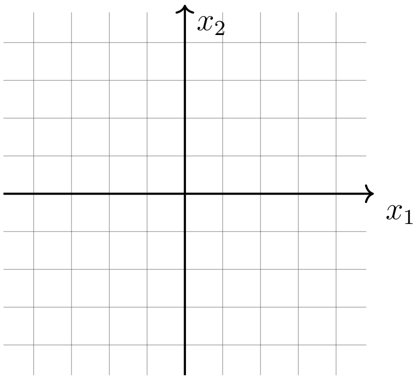
$$x_2 - 2x_1 = 4 \tag{5}$$

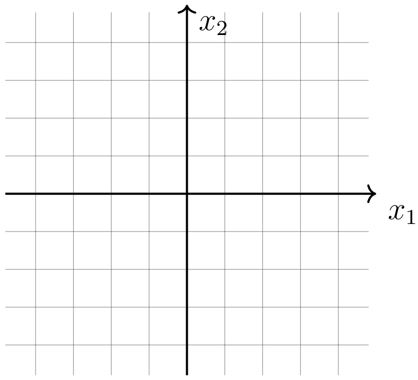
Each linear equation with two variables forms a line:



Now the intersection point of these two lines (,) is a solution.

With the above representation, we can easily see that two lines (equations) might be parallel (no solution), or might superpose (infinite solutions):





Thus, a system of linear equations has (i) no solution, or (ii) one solution (iii) infinite solutions.

In Example 1.5, we have already seen that for a linear system with only two unknowns, plotting the corresponding lines can give us the solution. Unfortunately, solving a larger (with more unknown variables) linear system is not always an easy task. However, there are some strategies to transform a complicated linear system to an equivalent (i.e. one with exactly same solution set) simpler one. For instance in Example 1.5, by subtracting (4) from (5) we can readily obtain x_1. A systematic way to solve systems of linear equations step by step is called "**Gaussian Elimination**". Before detailing this algorithm, let's see how a set of linear equations can be solved by simple algebraic operations:

Example 1.6.

$$x_1 - 3x_2 + x_3 = 4 \tag{6}$$

$$2x_1 - 8x_2 + 8x_3 = -2 \tag{7}$$

$$-6x_1 + 3x_2 - 15x_3 = 9 \tag{8}$$

To solve x_1, x_2, x_3 we must eliminate some unknowns from the equations. Let's try to remove x_1 from 7. To do that, we can add -2 times equation 6 to equation 7:

$$x_1 - 3x_2 + x_3 = 4 \tag{9}$$

$$0x_1 - 2x_2 + 6x_3 = -10 \tag{10}$$

$$-6x_1 + 3x_2 - 15x_3 = 9 \tag{11}$$

Similarly we can add 6 times equation 9 to the equation 11, to eliminate x_1 from the equation 11:

$$x_1 - 3x_2 + x_3 = 4 \tag{12}$$

$$0x_1 - 2x_2 + 6x_3 = -10 \tag{13}$$

$$0x_1 + -15x_2 - 9x_3 = 33 \tag{14}$$

We can simplify equations 13 and 14 by multiplying both sides with $\frac{1}{2}$ and $\frac{1}{3}$, respectively:

$$x_1 - 3x_2 + x_3 = 4 \tag{15}$$

$$0x_1 - 1x_2 + 3x_3 = -5 \tag{16}$$

$$0x_1 + -5x_2 - 3x_3 = 11 \tag{17}$$

Finally, in order to eliminate x_2 17, we can add -5 times equation 16 to the equation 17:

$$x_1 - 3x_2 + x_3 = 4 \tag{18}$$

$$0x_1 - 1x_2 + 3x_3 = -5 \tag{19}$$

$$0x_1 + 0x_2 - 18x_3 = 36 \tag{20}$$

Now we can easily solve equation 20 with only one unknown which is $x_3 = -2$. Plugging this solution to equation 19 yields $x_2 = -1$. Finally, the last unknown can be achieved by plugging known values for x_2 and x_3 into the equation 18, which gives: $x_1 = 3$.

The above operations can be performed in a more compact form with a **Matrix notation**. Let's look at another Example:

Example 1.7.

$$1x_1 + 2x_2 + 3x_3 = 2 \tag{21}$$

$$1x_1 + 1x_2 + 1x_3 = 2 \tag{22}$$

$$3x_1 + 3x_2 + 1x_3 = 0 \tag{23}$$

By identifying rows and columns, one can write the coefficients on the left hand side in a matrix form:

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 3 & 3 & 1 \end{pmatrix} \tag{24}$$

Which is called the **coefficient matrix**. By concatenating the right hand side of the linear set (as a column) to the right of this matrix, we obtain the **augmented matrix**:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 1 & 1 & 1 & 2 \\ 3 & 3 & 1 & 0 \end{array}\right) \tag{25}$$

The compact form of a set of linear equations ease the task of solving a set of linear equations. The system can be written in the form of $Ax = b$, where A is the coefficient matrix, b is the right hand side vector and x is the unknown vector. Now let's solve the above set of linear equations in a matrix form:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 1 & 1 & 1 & 2 \\ 3 & 3 & 1 & 0 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

Threfore, the solution is:

$$\left(\begin{array}{c} \\ \\ \\ \end{array}\right)$$

Let's attack another example, in which we need to do a Row Interchange:

Example 1.8.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 1 & 2 & 1 & 2 \\ 3 & 3 & 1 & 0 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

$$\begin{pmatrix} \\ \\ \\ \end{pmatrix}$$

As previously seen, a system of linear equations might not have a solution (inconsistent):

Example 1.9.

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 2 \\ 0 & -1 & -2 & 0 \\ 0 & -3 & -6 & 6 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} & & & \\ & & & \\ & & & \end{array}\right)$$

Definition 1.

- *A linear system is called **consistent** if ...*
- *A linear system is **inconsistent** if ...*

In general, "*Gaussian elimination* comprises two principle steps:
(a)
(b)
In doing so, we benefit from three elementary operations:
(i)
(ii)
(iii)
Please note that none of these row operations change the solution set of the linear system.

Definition 2. *Row Echelon Form (REF):*

- *all rows with at least one nonzero are above any rows of all zeros,*
- *reading from left to right, the first non-zero entry in any row (called **leading entry**) is in a column strictly to the right of the leading entry in the row above.*

Row Echelon Form (RREF), if additionally we have:

- *all pivots are equal to one,*
- *any column with a leading entry has zeros <u>above</u> and below it.*

Example 1.10.

$$2x_1 + 4x_2 + 4x_3 + 6x_4 = 0$$
$$x_1 + 2x_2 + 3x_3 + 4x_4 = 1$$
$$x_1 + 2x_2 + x_4 = -2$$

$$Ax = b$$

$$\begin{pmatrix} \quad & \quad & \quad & \quad \\ & & & \\ & & & \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \quad \\ \\ \end{pmatrix}$$

$$\left(\begin{array}{cccc|c} \quad & \quad & \quad & \quad & \quad \\ & & & & \\ & & & & \end{array}\right)$$

$$\left(\begin{array}{c|c} \quad & \quad \\ & \\ & \end{array}\right)$$

So the Row Echelon Form (REF) is:

$$\left(\begin{array}{c|c} \quad & \quad \\ & \\ & \end{array}\right)$$

$$\left(\begin{array}{c|c} \quad & \quad \\ & \\ & \end{array}\right)$$

And the Reduced Row Echelon (RREF) Form:

$$\left(\begin{array}{c|c} \quad & \quad \\ & \\ & \end{array}\right)$$

$$\begin{aligned} m &= \text{the number of equations (rows)} \\ n &= \text{the number of uknowns (columns)} \\ r &= \text{the number of pivots} \\ n - r &= \text{the number of } \textbf{free variables} \end{aligned}$$

If $n - r > 0$, we write the pivot variables in terms of free variables as parameters:

$$x = \begin{pmatrix} \quad \\ \\ \\ \end{pmatrix}$$

With arbitrary values for the free variables.

Example 1.11. *4 equations, 3 unknowns: Find the general solution of the linear system:*

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 4 & 3 & 0 \\ 6 & 4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \\ 4 \\ 9 \end{pmatrix}$$

Theorem 1. *Existence and Uniqueness of reduced row echelon form (RREF): every matrix A can be transformed into a matrix R in reduced row echelon form by using the elementary row operations. Matrix R is unique. In other words, it does not depend on the order in which the row operations are applied.*

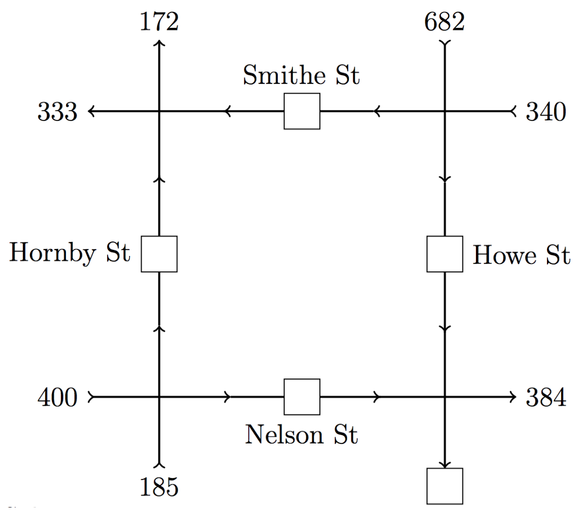
In short, we write a system of linear equations Ax = n in the form of augmented matrix. Next:

1. the forward elimination returns the REF. If the system is inconsistent, then we conclude that there is no solution and stop. Otherwise,

2. backward elimination returns the RREF $(R|d)$. We identify pivot columns and free columns in R to derive a general expression for all solutions x, which still contains the free variables as parameters.

Now let's concentrate on a more practical examples:

Example 1.12. *Transportation:*



Counts of vehicles per hour were collected at various locations along four one-way streets in downtown Vancouver. Assuming that there is no parking available, how many cars have passed the marked locations where no traffic counts were undertaken?

$$x = \begin{pmatrix} \\ \\ \\ \\ \end{pmatrix} + x_4 \begin{pmatrix} \\ \\ \\ \\ \end{pmatrix}$$

From a physical point of view, a solution is sensible only if

-

Thus, we only consider those solutions which satisfy

$$\left\{ \quad \right\} \longrightarrow \left\{ \quad \right\} \longrightarrow \left\{ \qquad \right\}$$

2 Sets of vectors

A n-tuple of numbers is called a vector. It can also be considered as a matrix with only one column. Vectors have many useful properties which make them a popular form of mathematical structure applicable to a wide range of real-life problems. We start with some simple definitions:

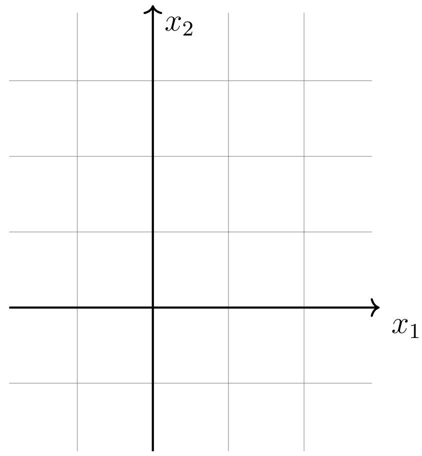
Definition 3.

- *When the vector a belongs to a space $\mathbb{R}^m$ it means that all m entries of a belong to Real numbers,*
- *Additive closure: $a_1 + a_2 \in V$ (Adding two vectors give a vector),*
- *Additive commutativity: $a_1 + a_2 = a_2 + a_1$. (Order of addition does not matter),*
- *Distributivity : $c(a_1 + a_2)$ (Scalar multiplication distributes over addition of vectors),*
- *Associativity: $c(a_1 \cdot a_2)$*

$\{a_1, a_2, a_3, ..., a_n\} \in \mathbb{R}^m$ is called a set of vectors in Real numbers where the order does not matter. One question that we need to answer is:

- Given $b \in \mathbb{R}^m$, can it be represented as a linear combination of $\{a_1, ..., a_n\}$?

Example 2.1. *Verify if $b = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$ can be expressed by a linear combination of $a_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $a_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. Find the coefficients of the linear combination.*



Definition 4. *Vector form for a linear set:*
A linear combination of equations can be viewed as a sum of basis vectors with unknown coefficients:

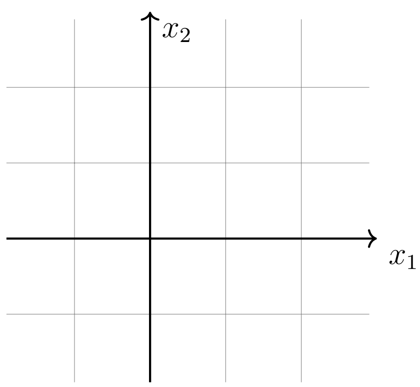
$$a_{11}x_1 + a_{12}x_2 = b_1 \qquad \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} x_1 + \begin{pmatrix} a_{21} \\ a_{22} \end{pmatrix} x_2 = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$
$$a_{21}x_2 + a_{22}x_2 = b_2$$

Definition 5. *For any set of* $\{a_1, ..., a_n\}$ *in* $\mathbb{R}^m$*, the set of all linear combinations of* $\{a_1, ..., a_n\}$ *is called the span of this vector set:*

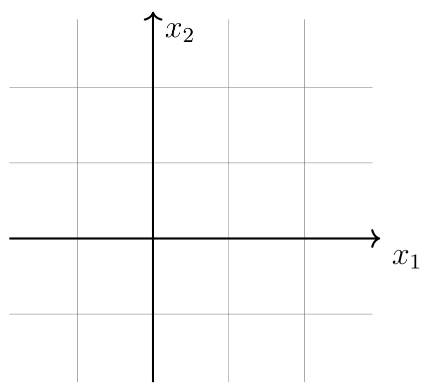
$$\textit{Span}\{a_1, ..., a_n\} := \{x_1 a_1 + ... + x_n a_n | x_1, ..., x_n \in \mathbb{R}\}$$

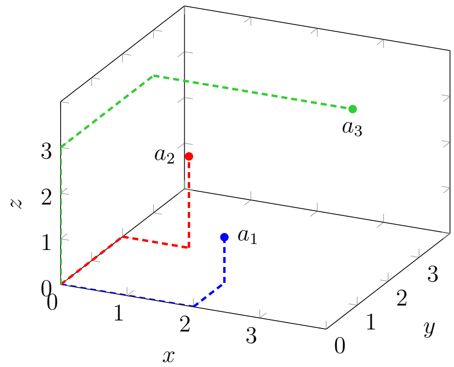
Example 2.2. *Verify if the following sets span the given space ?*

$a_1 = (1, 1), a_2 = (2, 2),$
$a_3 = (-1, -1)$



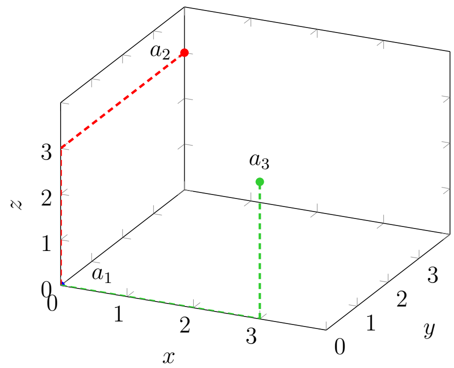
$a_1 = (-1, 2), a_2 = (1, 1)$





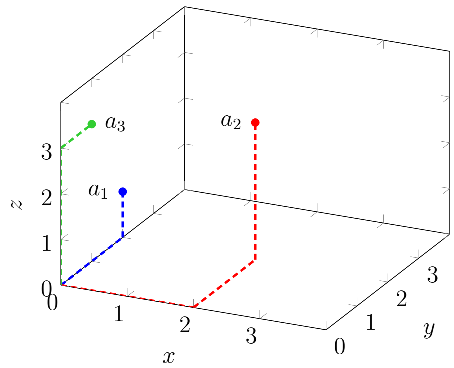
$a_1(2,1,1), a_2(1,2,2), a_3(3,3,3)$

- span$(a_1, a_2, a_3) \stackrel{?}{=} \mathbb{R}^3$



$a_1(0,0,0), a_2(0,4,3), a_3(3,0,3)$

- span$(a_1, a_2, a_3) \stackrel{?}{=} \mathbb{R}^3$



$a_1(0,2,1), a_2(2,2,3), a_3(0,1,3)$

- span$(a_1, a_2, a_3) \stackrel{?}{=} \mathbb{R}^3$

Example 2.3. *For what values of h will b be in Span$\{a_1, a_2, a_3\}$?*

$$a_1 = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 5 \\ -4 \\ -7 \end{pmatrix}, \quad a_3 = \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} -4 \\ 3 \\ h \end{pmatrix}$$

Definition 6. *Matrix form $Ax = b$ for a linear set:*
A linear combination of equations can be viewed as a product of coefficient matrix A and vector of unknowns x:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_2 + a_{22}x_2 &= b_2 \end{aligned} \qquad \begin{aligned} &[a_1 \;\; a_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = b \\ &A.x = b \end{aligned} \qquad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

How to compute $A.x = b$ more efficiently ?

Theorem 2. *if A is a $m \times n$ matrix, u and v are vectors in $\mathbb{R}^n$, and c is a scalar, then:*

- $A(u + v) = Au + Av$
- $A(cv) = c(Au)$

Theorem 3. *The following statements are equivalent:*

- *span*$(a_1, a_2, ..., a_n) = \mathbb{R}^m$.
- *For any vector $b \in \mathbb{R}^m$ there exist numbers $x_1, .., x_2$ such that:*

$$x_1 a_1 + ... + x_n a_n = b.$$

- *For any vector $b \in \mathbb{R}^m$ the problem $Ax = b$ has at least one solution $x \in \mathbb{R}^n$.*
- *The matrix A has m pivots, one pivot in each row.*

Example 2.4. *Does $\{a_1, a_2, a_3\}$ span the $\mathbb{R}^3$?*

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 3 \\ -2 \\ 2 \end{pmatrix}, \quad a_3 = \begin{pmatrix} -4 \\ 6 \\ -1 \end{pmatrix}$$

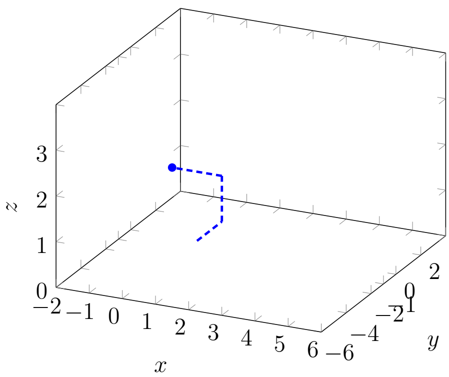
3 Solution sets of linear systems

3.1 Homogeneous system

Any linear system in the form of $Ax = 0$ is called a homogeneous system. There is always *at least* one solution for any homogeneous system, that is: $x = 0$. This solution is called the **trivial** solution. Any other non-zero vector that satisfies the linear system is called a **non-trivial** solution.

Example 3.1.

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 4 & 3 & 0 \\ 6 & 4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

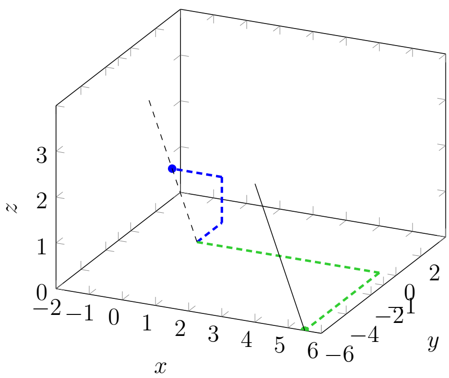


Hence, a homogeneous system has a non-trivial solution if and only if there is at least one free variable in the system.

3.2 Nonhomogeneous system

Example 3.2.

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 4 & 3 & 0 \\ 6 & 4 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \\ 4 \\ 9 \end{pmatrix}$$



Summery: Writing the solution of a linear system in parametric form can be achieved by following steps:

- Row reduction (Forward/backward elimination)
- Express pivot variables in terms of free variables (put free variables on the right hand side)
- Write the solution vector x in terms of free variables, if any.
- Decompose x into a linear combination of vectors using the free variables as parameters.

Example 3.3. *A muesli company is planning to introduce a new product. The new muesli mix will be composed of rolled oats, raisins, almonds, dried blueberries and banana chips, for which the following nutritional values are known:*

nutrition per 100 gr	*Rolled oats*	*Raisins*	*Almonds*	*Dried blueberries*	*Banana chips*
Carbohydrates	*70 gr*	*80 gr*	*20 gr*	*90 gr*	*60 gr*
Fat	*6 gr*	*1 gr*	*50 gr*	*2 gr*	*35 gr*
Protein	*15 gr*	*3 gr*	*21 gr*	*3 gr*	*2 gr*
Other	*9 gr*	*16 gr*	*9 gr*	*5 gr*	*3 gr*

In what proportion the nutrients should the ingredients be combined to achieve a nutrietion profile of carbohydrates: fat : protein : other $= 6 : 1 : 2 : 1$

From a physical viewpoint, a solution is only sensible if:

-

Therefore,

Conclusion: To solve a problem that stems from an application of linear algebra, we first identify equations and unknowns to set up a system of linear equations that models this problem. Once we have found all mathematical solutions of this linear system by Gaussian elimination, we interpret these solutions from the perspective of the application. It is important to note that mathematically correct answers may not always be meaningful in real life.

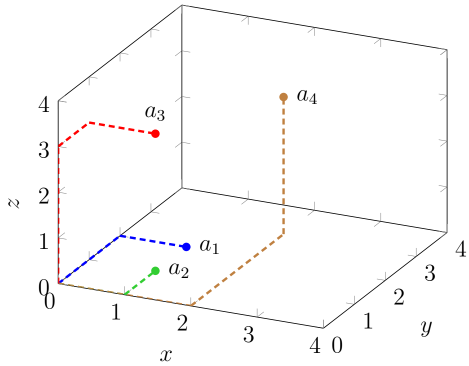
4 Linear independence

Let $\{a_1, a_2, ..., a_n\} \in \mathbb{R}^m$. Can we write any of vectors $a_1, a_2, ..., a_n$ in the form of a linear combination of other vectors?

Example 4.1.

$$a_1 = [1, 2, 0],$$
$$a_2 = [1, 1, 0],$$
$$a_3 = [1, 1, 3]$$

$$a_1 = [1, 2, 0],$$
$$a_3 = [1, 1, 3],$$
$$a_4 = [2, 3, 3]$$



Definition 7. *A set of vectors $a_1, a_2, ..., a_n$ is called linearly independent if the equation*

$$a_1x_1 + a_2x_2 + ... + x_na_n = 0$$

has only the trivial solution.

The following statements are equal:

- The family $(a_1, a_2, ..., a_n) = \mathbb{R}^m$ is linearly independent.
- The problem
$$x_1a_1 + x_2a_2 + ... + x_na_n = 0$$
has only one solution which is $x_1 = x_2 = x_3 = ... = x_n = 0$.
- The problem $Ax = 0$ has only the trivial solution $x = 0 \in \mathbb{R}^n$.
- The matrix A has n pivots, one pivot in each column.

These ideas can be better pictured in the following example.

Example 4.2. *Determine if the following families of vectors span the full space $\mathbb{R}^m$, if they are linearly independent?*

$$a_1 = \begin{pmatrix}1\\1\\3\end{pmatrix} \quad a_2 = \begin{pmatrix}2\\1\\3\end{pmatrix} \quad a_3 = \begin{pmatrix}3\\1\\1\end{pmatrix}$$ (*c.f.* Example 1.7)

$$a_1 = \begin{pmatrix}1\\1\\3\end{pmatrix} \quad a_2 = \begin{pmatrix}2\\2\\3\end{pmatrix} \quad a_3 = \begin{pmatrix}3\\1\\1\end{pmatrix}$$ (*c.f.* Example 1.8)

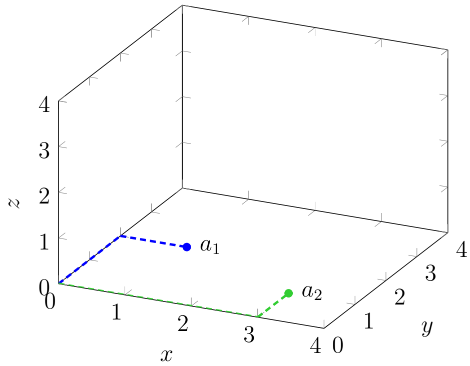
$$a_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix} \quad a_2 = \begin{pmatrix}2\\-1\\0\\0\end{pmatrix} \quad a_3 = \begin{pmatrix}3\\-2\\-6\\5\end{pmatrix} \quad a_4 = \begin{pmatrix}3\\-2\\-6\\5\end{pmatrix}$$

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad a_2 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix} \quad a_3 = \begin{pmatrix} 3 \\ -2 \\ -6 \\ 5 \end{pmatrix} \quad a_4 = \begin{pmatrix} 3 \\ -2 \\ -6 \\ 5 \end{pmatrix} \quad a_5 = \begin{pmatrix} 1 \\ 2 \\ 8 \\ 2 \end{pmatrix}$$

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad a_2 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix} \quad a_3 = \begin{pmatrix} 3 \\ -2 \\ -6 \\ 5 \end{pmatrix}$$

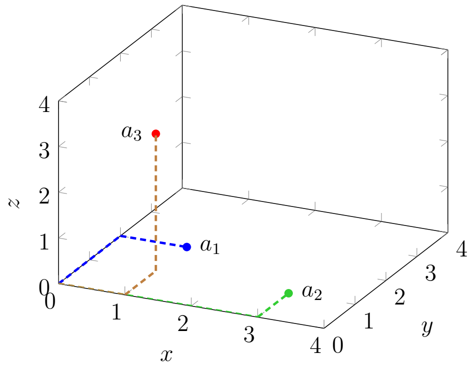
Example 4.3.

$$a_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad a_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$$



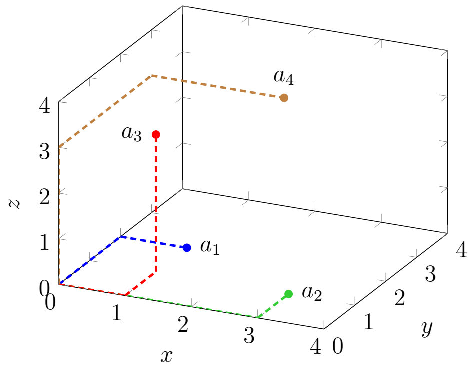
$$a_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad a_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$$

$$a_3 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$$



$$a_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad a_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$$

$$a_3 = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \quad a_4 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}$$

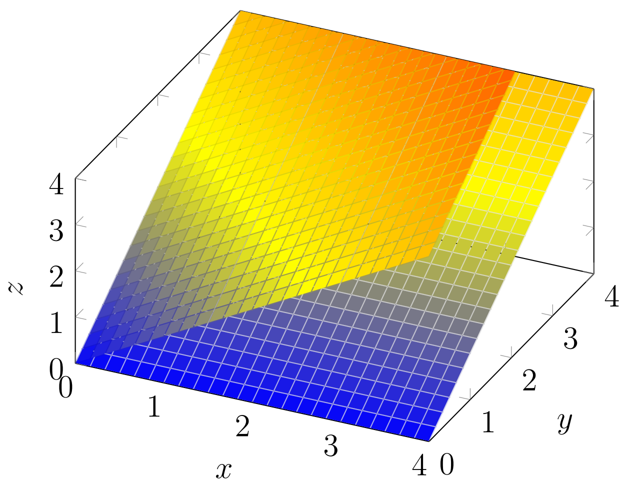


Sets of Vectors	Linear Systems	Pivots
The vectors $a_1, \ldots, a_n \in \mathbb{R}^m$ *span the whole space:* $\text{span}(a_1, \ldots, a_n) = \mathbb{R}^m$.	*Existance of solutions:* for any $b \in \mathbb{R}^m$ the problem $Ax = b$ has *at least* one solution.	The matrix $A \in \mathbb{R}^{m \times n}$ has a pivot in each of its m rows. In particular we must have $m \leq n$
The vectors $a_1, \ldots, a_n \in \mathbb{R}^m$ *are linearly independent.*	*Uniqueness of the solutions:* for any $b \in \mathbb{R}^m$ the problem $Ax = b$ has *at most* one solution.	The matrix $A \in \mathbb{R}^{m \times n}$ has a pivot in each of its n columns. In particular we must have $m \geq n$
The vectors $a_1, \ldots, a_n \in \mathbb{R}^m$ *form a basis* of $\mathbb{R}^m$.	*Existance and uniqueness of solutions:* for any $b \in \mathbb{R}^m$ the problem $Ax = b$ has *exactly* one solution.	The matrix $A \in \mathbb{R}^{m \times n}$ has a pivot in each of its m rows and n columns. In particular, we must have $m = n$

Reminder: To find the number of pivots in A, we use Forward elimination to transform A to REF.

Example 4.4. *The two homogeneous equations, below, define two planes through the origin in $\mathbb{R}^3$. Find a parametric vector form for the line of intersection of the two planes.*

$$
\begin{aligned}
y - z &= 0 \\
-x - y + z &= 0
\end{aligned}
$$



5 Linear transformation

In this sections we are considering functions between vector space $\mathbb{R}^m$ and $\mathbb{R}^n$.
Example 5.1.

$\mathbb{R} \longrightarrow \mathbb{R}$ $\qquad$ $\mathbb{R}^2 \longrightarrow \mathbb{R}^3$ $\qquad$ $\mathbb{R}^3 \longrightarrow \mathbb{R}^1$

Definition 8.
- *(Domain) The set of all vectors x for which $T(x)$ is defined*
- *(Range) The set of all vectors of the form $T(x)$ for some x in the domain of T.*
- *(Codomain) The set that contains the range of T.*

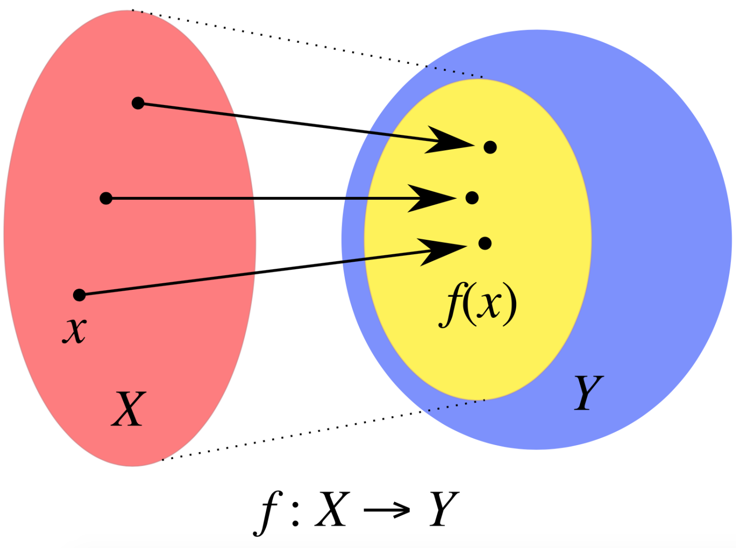


Figure 1: X: Domain, Y: Codomain, $f(x)$: range

However we will only be looking at functions with a special property that we refer to as *linearity*. As it turns out, such linear maps provide yet another interpretation of matrices multiplied with vectors.

Definition 9. *A function $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is said to be a linear map or a linear transformation if it satisfies the following two properties:*

$$\text{(Additivity)} \quad T(u+v) = T(u) + T(v) \qquad \text{for any } u, v \text{ in } \mathbb{R}^n$$

$$\text{(Homogeneity)} \quad T(cu) = cT(u) \qquad \text{for any } u, v \text{ in } \mathbb{R}^n \text{ and } c \text{ in } \mathbb{R}$$

These two properties can be simply combined and represented by:

$$T(cu + cv) = cT(u) + cT(v) \qquad \textit{for any } u, v \textit{ in } \mathbb{R}^n \textit{ and } c \in \mathbb{R}$$

Example 5.2. *Check if T is linear:*

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2 \qquad T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_2 \\ 3x_1 \end{bmatrix}$$

Example 5.3. *Check if T is linear:*

$$T : \mathbb{R}^1 \mapsto \mathbb{R}^1 \qquad T\begin{bmatrix} x_1 \end{bmatrix} = \begin{bmatrix} 2x_1 - 1 \end{bmatrix}$$

5.1 Matrix of linear transformation

Up to now, we have looked at linear transformation as a *formula*. In the following we show that linear transformation is simply another interpretation of matrices multiplied by vectors. In particular, a linear transformation can be seen as matrix A that "acts" on the vector x and produces the vector b:

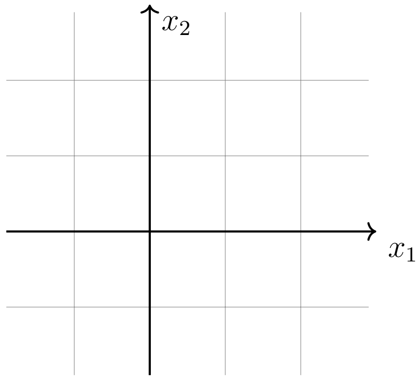
$$A.x = b$$

Hence, many familiar properties of the linear systems in matrix form can be related to the linear transformations.

Example 5.4. *(dilation or contraction)*

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

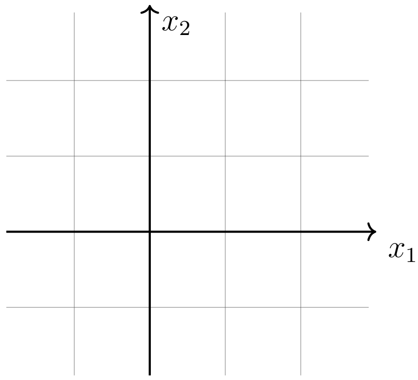
$$A = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}, \quad c \in \mathbb{R}$$



Example 5.5. *(rotation by $90°$)*

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

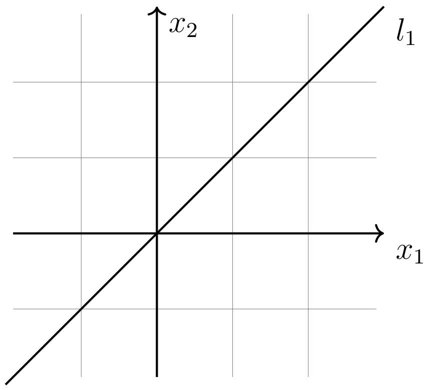
$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$



Example 5.6. *(reflection across the line $l_1 : x_2 = x_1$)*

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

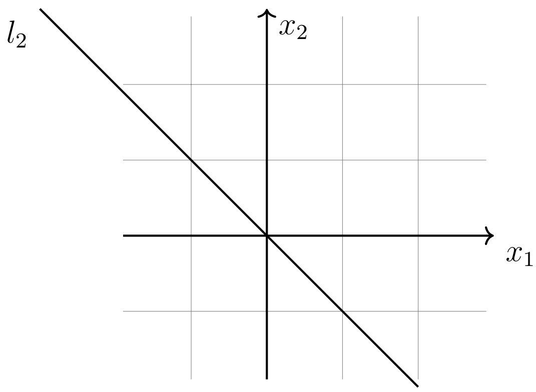
$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$



Example 5.7. *(reflection across the line $l_2 : x_2 = -x_1$)*

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

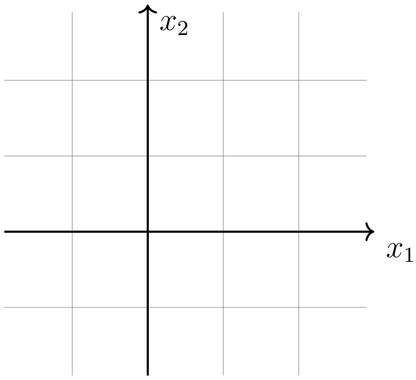
$$A = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$



Example 5.8. *(reflection across the origin)*

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

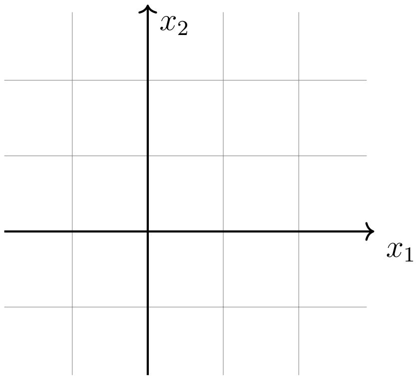
$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$



Example 5.9. *(Projection on the axis x_2)*

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

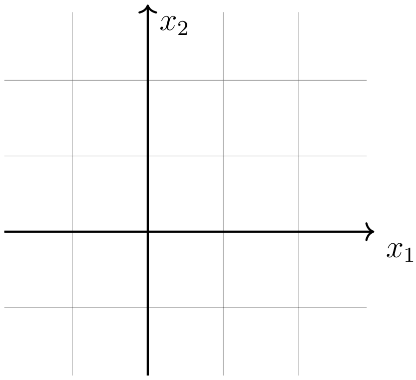
$$A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$



Example 5.10. *(Horizontal expansion)*

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

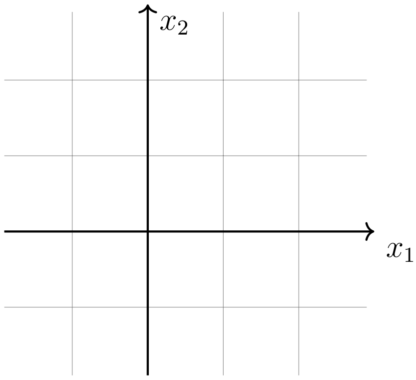
$$A = \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix}, \quad c \in \mathbb{R}$$



Example 5.11. *(Horizontal shear)*

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

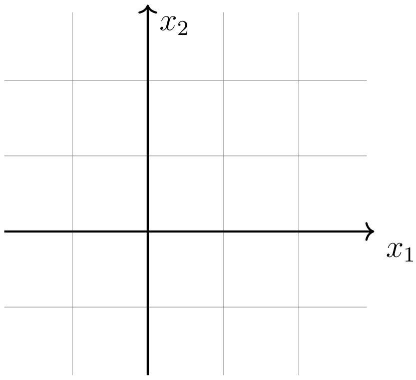
$$A = \begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix}, \quad c \in \mathbb{R}$$



Example 5.12. *(Vertical shear)*

$$T : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

$$A = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}, \quad c \in \mathbb{R}$$



Example 5.13. *Imagine T is a linear transformation* $\mathbb{R}^2 \mapsto \mathbb{R}^3$*. For given input unit vectors* $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *(these vectors are columns of the identity matrix* $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$*), the outputs are the followings:*

$$T(e_1) = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \qquad\qquad T(e_2) = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix},$$

Find the standard matrix of T.

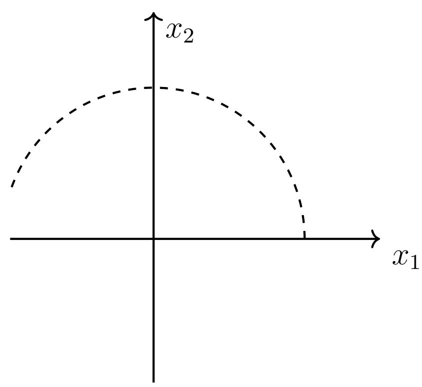
Theorem 4. *Let* $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ *be a linear transformation. There exists a unique matrix* A *such that:*

$$T(x) = Ax, \qquad\qquad \textit{for all } x \textit{ in } \mathbb{R}^n$$

Matrix A *is a* $m \times n$ *matrix whose* j*th column is the vector* $T(e_j)$*, where* e_j *is the* jth *column of the identity matrix in* $\mathbb{R}^n$*:*

$$A = [T(e_1) \qquad \ldots \qquad T(e_n)]$$

Example 5.14. *Find the standard matrix* A *for the transformation* T” $\mathbb{R}^2 \mapsto \mathbb{R}^2$ *which rotates all the inputs with the angle* ϕ



Example 5.15. *Let $T : \mathbb{R}^2 \mapsto \mathbb{R}^3$ be a linear transformation with following formula:*

$$T(x_1, x_2) = \begin{bmatrix} x_1 + 2x_2 \\ 2x_1 + x_2 \\ 0 \end{bmatrix}$$

First, identify the standard matrix of transformation. Second, identify the domain, codomain and range of the transformation. Is any arbitrarily chosen vector in $\mathbb{R}^3$ an image of at least one x ?

Theorem 5. *A mapping $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is said to be* ***onto*** *$\mathbb{R}^m$ if each $b \in \mathbb{R}^m$ is the image of at least one $x \in \mathbb{R}^n$. This is true if and only if the columns of standard matrix A* ***span*** *$\mathbb{R}^m$*

Example 5.16. *In Example 5.15, for an arbitrarily chosen vector $b \in \mathbb{R}^3$, how many input vectors $x \in \mathbb{R}^2$ exist for which $T(x) = b$?*

- if b is outside the range of $T(x)$...
- if b is inside the range of $T(x)$...

Theorem 6. *A mapping $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is said to be* ***one-to-one*** *if each $b \in \mathbb{R}^m$ is the image of at most one $x \in \mathbb{R}^n$. This is true if and only if the columns of standard matrix A are linearly independent*

Example 5.17. *Consider the following linear transformation $T : \mathbb{R}^3 \mapsto \mathbb{R}^2$:*

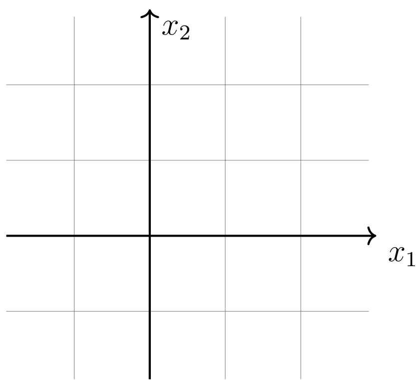
$$T(x_1, x_2, x_3) = \begin{bmatrix} x_1 + 2x_2 \\ x_1 + x_2 + 2x_3 \end{bmatrix}$$

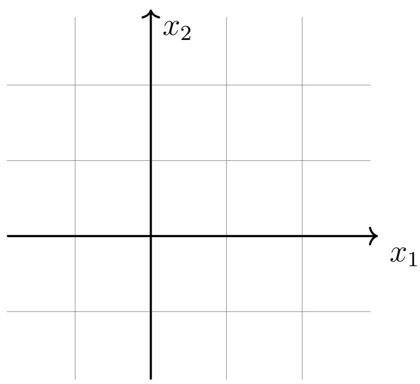
Is T one-to-one ? Is T onto $\mathbb{R}^m$?

Linear Maps	Sets of Vectors	Linear Systems	Pivots
$T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is *onto* $\mathbb{R}^m$.	Vectors $a_1, ..., a_n \in \mathbb{R}^m$ span $\mathbb{R}^m$.	For any $b \in \mathbb{R}^m$ the problem $Ax = b$ has *at least* one solution.	The matrix $A \in \mathbb{R}^{m \times n}$ has a pivot in each of its m rows. In particular we must have $m \leq n$
$T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is *one-to-one*.	Vectors $a_1, ..., a_n \in \mathbb{R}^m$ *are linearly independent.*	For any $b \in \mathbb{R}^m$ the problem $Ax = b$ has *at most* one solution.	The matrix $A \in \mathbb{R}^{m \times n}$ has a pivot in each of its n columns. In particular we must have $m \geq n$
$T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is *onto* $\mathbb{R}^m$ and *one-to-one*.	Vectors $a_1, ..., a_n \in \mathbb{R}^m$ *form a basis* of $\mathbb{R}^m$.	For any $b \in \mathbb{R}^m$ the problem $Ax = b$ has *exactly* one solution.	The matrix $A \in \mathbb{R}^{m \times n}$ has a pivot in each of its m rows and n columns. In particular, we must have $m = n$

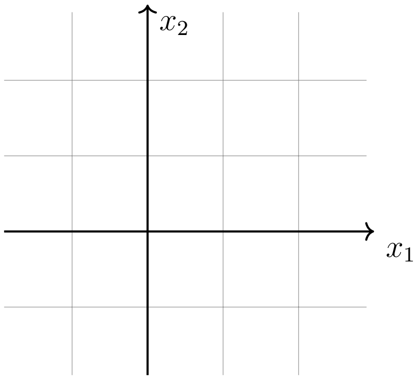
Reminder: To find the number of pivots in A, we use Forward elimination to transform A to REF.

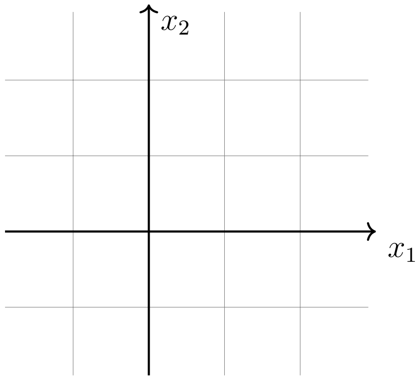
Example 5.18. *Find the standard matrix for a linear transformation $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ with the following features: it first performs a horizontal shear: it maps e_2 to $e_2 - 2e_1$ (and it leaves e_1 unchanged). Then it reflects the results through the origin. Is this linear map onto $\mathbb{R}^2$? Is it one-to-one ?*





Example 5.19. *Find the standard matrix for a linear transformation $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ with the following features: it first rotates the points (about the origin) through $-\pi/2$. Next, it projects every point onto the x_2 axis ? Is this linear map onto $\mathbb{R}^2$? Is it one-to-one ?*





6 Matrix Algebra

In linear algebra, we encounter three basic types of arithmetic operations that involves scalars $\lambda \in \mathbb{R}$ and matrices:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n} \qquad B = \begin{pmatrix} b_{11} & \cdots & b_{1q} \\ \vdots & \ddots & \vdots \\ b_{p1} & \cdots & b_{pq} \end{pmatrix} \in \mathbb{R}^{p \times q}$$

- Multiplication with scalars :

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}$$

- Matrix addition $\longrightarrow$ only works if A and B have the same size:

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1q} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{p1} & \cdots & a_{mn} + b_{pq} \end{pmatrix}$$

- Matrix multiplication $\longrightarrow$ only works if $n = p$ (number of columns in A is equal to number of rows in B):

$$AB = \begin{pmatrix} [a_{11} \cdots a_{1n}] \begin{bmatrix} b_{11} \\ \vdots \\ b_{p1} \end{bmatrix} & \cdots & [a_{11} \cdots a_{1n}] \begin{bmatrix} b_{1q} \\ \vdots \\ b_{pq} \end{bmatrix} \\ \vdots & \ddots & \vdots \\ (a_{m1} \cdots a_{mn}) \begin{bmatrix} b_{11} \\ \vdots \\ b_{p1} \end{bmatrix} & \cdots & [a_{m1} \cdots a_{mn}] \begin{bmatrix} b_{1q} \\ \vdots \\ b_{pq} \end{bmatrix} \end{pmatrix}$$

Example 6.1. *Compute the matrix multiplication when possible.*

- $AB = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$

- $AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$

- $AB = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \begin{pmatrix} 7 & 8 \end{pmatrix}$

- $AB = \begin{pmatrix} 7 & 8 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$

If A is a $m \times n$ matrix, and if B is a $n \times p$ matrix with columns $b_1, \cdots b_p$, another way to compute the multiplication is to write AB as a linear combination of columns of matrix B:

$$AB = A[b_1 \quad b_2 \quad \cdots \quad b_p] = [Ab_1 \quad Ab_2 \quad \cdots \quad Ab_p]$$

Example 6.2. *Compute AB where $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 1 & 1 \end{pmatrix}$*

Properties of matrix multiplication: Let A be a $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

- $A(BC) = (AB)C$ Associativity
- $A(B + C) = AB + AC$ Distributivity
- $(B + C)A = BA + CA$
- $r(AB) = (rA)B = A(rB)$,
- $I_m A = A = AI_n$, I_n is a $n \times n$ identity matrix
- $A^k = \underbrace{A \cdots A}_{k}$

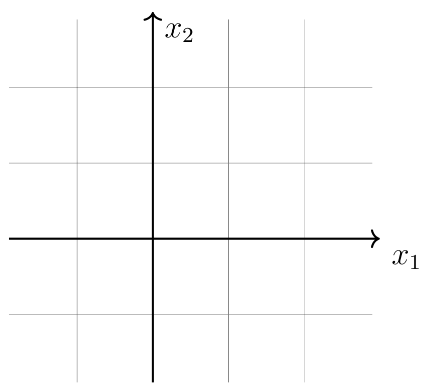
Example 6.3. *Given* $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ *and* $B = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$, *compute* AB *and* BA *and verify if* $AB = BA$.

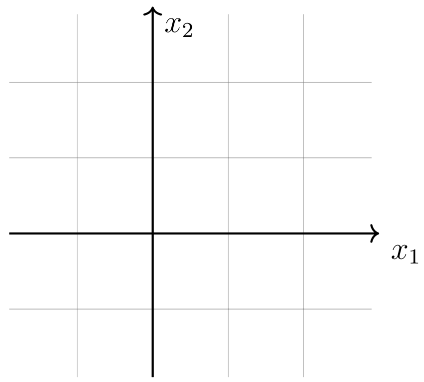
Warning*: in matrix multiplication order matters !*

Example 6.4. *Let* $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ *and* $P : \mathbb{R}^2 \mapsto \mathbb{R}^2$ *be the linear transformations corresponding to the rotation by* $\pi/2$ *and the projection on the axis* x_1. *What is the standard matrix corresponding to the composition of these two linear maps (first* T *and second* P*). Hint: use the associativity rule !*

$$T(x) = A_1 x \qquad P(x) = A_2 x$$

$$A_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$





Definition 10. *Given an $m \times n$ matrix A, the **transpose** of A is the $m \times n$ matrix, denoted by A^T, whose columns are formed the corresponding rows of A.*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

Example 6.5. *Find the transpose of the following matrices:*

- $$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, A^T =$$

- $$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}, A^T =$$

Properties of the transpose of a matrix:

- $(A^T)^T = A$

- $(A+B)^T = A^T + B^T$

- for any scalar r, $(rA)^T = rA^T$
- $(AB)^T = B^T A^T \rightarrow$ General Form: $(AB \cdots YZ)^T = Z^T Y^T \cdots B^T A^T$

7 Inverse of a matrix

This section addresses the question how we can undo the action of a matrix or linear map, provided that this is possible at all: if Ax give b, the $A^{-1}b$ should give x. Such an inverse matrix can only exist if:

- for any $b \in \mathbb{R}^m$ there is an $x \in \mathbb{R}^n$ such that $Ax = b$ (Existence of the solution)
- there are not two or more $x \in \mathbb{R}^n$ such that $Ax = b$. (Uniqueness of the solution)

Hence, we only consider square matrices $A \in \mathbb{R}^{n\times n}$ in this section and look for A^{-1} of the same size.

Definition 11. *If for a matrix $A \in \mathbb{R}^{n\times n}$ there exist a matrix $A^{-1} \in \mathbb{R}^{n\times n}$ such that :*

$$AA^{-1} = I_n \qquad \textit{and} \qquad A^{-1}A = I_n$$

then A is said to be invertible and A^{-1} is called the inverse of A. Otherwise A is singular.

Note that if $A \in \mathbb{R}^n$ is invertible, after row reduction it reduces to I_n (Why ?).
If A is invertible, how can we compute A^{-1} ? If $A \in \mathbb{R}^{2\times 2}$, this is very easy:

Theorem 7. *Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. If $ad - bc \neq 0$, then A is invertible and*

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Example 7.1. $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $A^{-1} =$

For an invertible matrix $A = [y_1 \quad y_2 \quad \cdots \quad y_j] \in \mathbb{R}^{n\times n}$:

$$\begin{aligned}
(1\text{th column of } A^{-1}) &= y_1 = A^{-1}e_1 &\longrightarrow\quad& Ay_1 = e_1\\
(j\text{th column of } A^{-1}) &= y_j = A^{-1}e_j &\longrightarrow\quad& Ay_j = e_j\\
(n\text{th column of } A^{-1}) &= y_n = A^{-1}e_n &\longrightarrow\quad& Ay_n = e_n
\end{aligned}$$

Therefore, we have to solve n simultaneous linear system:

$$(A \mid e_1 \; e_2 \; \cdots \; e_n) \longrightarrow (I_n \mid y_1 \; y_2 \; \cdots \; y_n)$$
$$(A^{-1} \mid I_n) \longrightarrow (I_n \mid A^{-1})$$

Example 7.2. *Determine whether or not*

$$A = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & -2 \\ 0 & 1 & 2 \end{pmatrix}$$

is invertible and if so, find its inverse matrix.

Properties of matrix inversion:

- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^{-1})^{-1} = A$
- $(A^T)^{-1} = (A^{-1})^T$

8 Characteristics of invertible matrices

We already know that matrix $A \in \mathbb{R}^{n\times n}$ is invertible if and only if the linear system $Ax = b$ has exactly one solution. This can be interpreted in terms of properties of linear transforamtion. A linear transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ is said to be invertible if there exist a function $S : \mathbb{R}^n \mapsto \mathbb{R}^n$ such that:

$$S(T(x)) = x \qquad \text{for all } x \in \mathbb{R}^n$$
$$T(S(x)) = x \qquad \text{for all } x \in \mathbb{R}^n$$

In that case, A is also invertible and $S(x) = A^{-1}x$.

Theorem 8. *A linear transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ is invertible, if and only if the corresponding standard matrix $A \in \mathbb{R}^n \mapsto \mathbb{R}^n$ is also invertible. In that case, the inverse of the linear transformation $S : \mathbb{R}^n \mapsto \mathbb{R}^n$ reads as $S(x) = A^{-1}x$*

Example 8.1. *Let $A \in R^{n\times n}$. Check if the following statements are true:*

- *if A is invertible, then its columns span $\mathbb{R}^n$.*
- *If A is invertible, then the corresponding linear transformation is one-to-one.*
- *if the columns of A are linearly independent, then A is invertible.*
- *if the equation $Ax = b$ is inconsistent for some $b \in \mathbb{R}^n$, then the equation $Ax = 0$ has only the trivial solution.*
- *If the first two columns of A are equal, A is not invertible.*
- *If the equation $Ax = 0$ has only the trivial solution, then A is row equivalent to the $n \times n$ identity matrix.*
- *Let $B \in \mathbb{R}^{n\times n}$. If AB is invertible, then A is invertible.*

9 Subspace and basis

We have already encountered various subsets of vector spaces, e.g.

-

-

-

Some subsets are especial because they form another vector space inside the larger vector space:

Definition 12. *A non-empty set $H \subset \mathbb{R}^n$ is said to be a subspace of $\mathbb{R}^n$ if it satisfies the following two properties:*

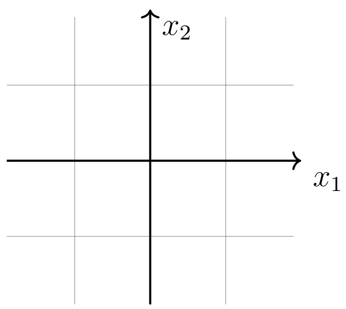
- *(Closedness under Addition)*
- *(Closedness under Scalar Multiplication)*

Or it can be combined to:

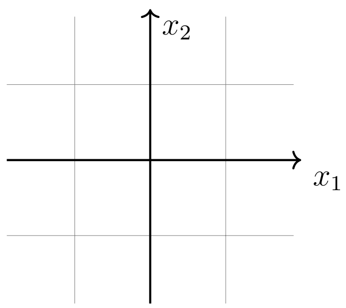
- *(Closedness under Linear Combination)*

Example 9.1. *Determine if the following subsets of $\mathbb{R}^2$ are subspaces?*

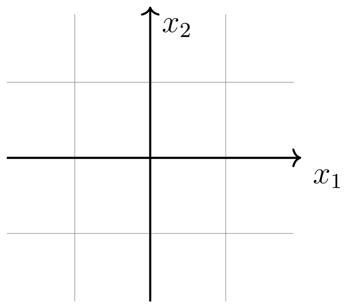
- $H = \{x \in \mathbb{R}^2 | x_2 = 2x_1\}$



- $H = \{x \in \mathbb{R}^2 | x_2 = x_1 - 1\}$



- $H = \{x \in \mathbb{R}^2 | x_2 x_1 \geq 0\}$



- $H =$ *Span*(u, v) $\qquad u, v \in \mathbb{R}^n$

Example 9.2. *If H is a subspace of $\mathbb{R}^3$, then H is either*

-
-
-
-

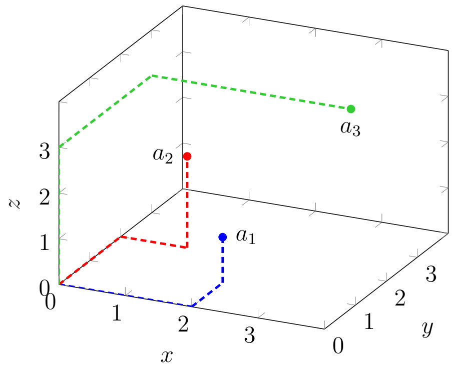
Definition 13. *A family of vectors $(v_1, ..., v_d) \subset \mathbb{R}^n$ is said to*

- *span the subspace $H \subset \mathbb{R}^n$ if*
- *be a basis for the subspace $H \subset \mathbb{R}^n$ if*

Definition 14. *(Dimension of a vector space) The dimension dim H of a vector space H is the number of vectors in every basis for H.*

Example 9.3. *Let*

$$H = span\left(\begin{bmatrix}2\\1\\1\end{bmatrix}, \begin{bmatrix}1\\2\\2\end{bmatrix}, \begin{bmatrix}3\\3\\3\end{bmatrix}\right)$$



Conclusions:

- A subspace is a vector space nested inside another vector space.

- To prove that a subset $H \subset \mathbb{R}^n$ is also a subspace we have to show that: 1) $H \neq \emptyset$ by giving an example of one vector in H (0 always works if H is actually a subspace) 2) if u and v are any two vectors in H, the all their linear combinations $\lambda u + \mu v$ must be contained in H as well.

  To prove that a subset $H \subset \mathbb{R}^n$ is *not* a subspace we need to find a counterexample that violates one of the conditions.

- $\mathbb{R}^n$ is the space of all vectors with n real-valued components. A $d-$dimensional subspace $H \subset \mathbb{R}^n$ is *isomorphic* to ("of the same shape as") the space of all vectors with only d components: $H \cong \mathbb{R}^d$.

10 Column space and Null space

In this section we are going to use vector-space language to describe general linear system $Ax = b$ and their solutions. Two subspaces associated with the matrix A:

- column space:

- null space: All solutions of $Ax = 0$

Definition 15. *Let $A \in \mathbb{R}^{m\times n}$ be the matrix with columns $a_1, \cdots a_n \in \mathbb{R}^m$.*

- *the set of all linear combinations of the columns of A is called the column space col(A)*

- *The set of all solutions of the homogeneous problem $Ax = 0$ is called the null space nul(A) or the kernel ker(A).*

Example 10.1. *Let $A \in \mathbb{R}^{m\times n}$. Show that col(A) and null(A) are subspace.*

Example 10.2. *For each of the following matrices (from examples 1.10 and 1.11), find a basis C for the column space and a basis N for the null space.*

$$A = \begin{pmatrix} 2 & 4 & 4 & 6 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 0 & 1 \end{pmatrix} \qquad A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 4 & 3 & 0 \\ 6 & 4 & 1 \end{pmatrix}$$

Solution:

$$\mathrm{RREF}(A) = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathrm{RREF}(A) = \begin{pmatrix} 1 & 0 & 3/2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

In Example 1.10 we found that for

$$b = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

all solutions of $Ax = b$ are

$$x = \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

In Example 1.10 we found that for

$$b = \begin{pmatrix} 5 \\ 10 \\ 4 \\ 9 \end{pmatrix}$$

all solutions of $Ax = b$ are

$$x = \begin{pmatrix} 11/2 \\ -6 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -3/2 \\ 2 \\ 1 \end{pmatrix}$$

Definition 16. *(Rank A) The rank of a matrix A, is the dimension of the col(A). That is equivalent to say that the rank of A is equal to the number of pivots in A.*

Theorem 9. *(Rank theorem) Let $A \in \mathbb{R}^{m\times n}$ (m rows and n columns) be a matrix of rank r.*

dim col$(A) =$ *dim null*$(A) =$

Conclusions: We assume that $A \in \mathbb{R}^{m\times n} \in$ has m rows and n columns and r pivots.

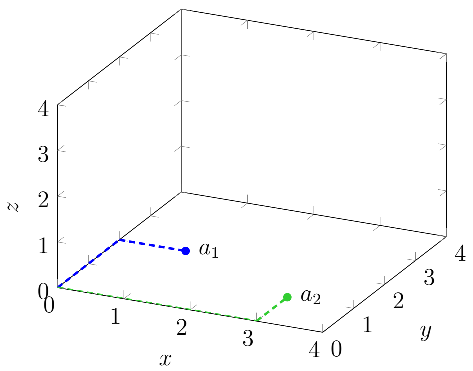
- The null space of A is a subspace of the input space $\mathbb{R}^n$. The column space is a subspace of the output space $\mathbb{R}^m$.
- The dimension of the subspace is the number of free columns in A, namely $n - r$. The dimension of the column space is the number of pivots in A, namely r.
- In order to find a basis for col(A) and null(A), we first do the row reduction and achieve the RREF. The solution of $Ax = 0$ in a parametric vector form readily provide us with a basis for null(A). The pivot columns of A (from the original matrix A, *not* the RREF of A) forms a basis for col(A).
- Important special cases are the smallest possible null space null$(A) = \{0\}$ and the largest possible column space col$(A) = \mathbb{R}^m$:

Linear Maps	Sets of Vectors	Linear Systems	Rank	Pivots
$T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is *onto* $\mathbb{R}^m$.	Vectors $a_1, ..., a_n \in \mathbb{R}^m$ span $\mathbb{R}^m$.	For any $b \in \mathbb{R}^m$ the problem $Ax = b$ has <u>*at least*</u> one solution.	A has full row rank $r = m$	A has a pivot in every row.
$T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is *one-to-one*.	Vectors $a_1, ..., a_n \in \mathbb{R}^m$ *are linearly independent.*	For any $b \in \mathbb{R}^m$ the problem $Ax = b$ has <u>*at most*</u> one solution.	A has full column rank $r = n$	A has a pivot in every column.
$T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is *onto* $\mathbb{R}^m$ and *one-to-one*.	Vectors $a_1, ..., a_n \in \mathbb{R}^m$ *form a basis* of $\mathbb{R}^m$.	For any $b \in \mathbb{R}^m$ the problem $Ax = b$ has <u>*exactly*</u> one solution.	A has full rank $r =$ $= m = n$	A has a pivot in every row and column.

Knowing about the two fundamental subspaces of A which are null(A) and col(A), we can know about the number of solutions of the system $Ax = b$:

Example 10.3. *Let $\{a_1, a_2\}$ be a basis for subspace H. Show that b is in the subspace H.*

$$a_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 3 \\ 0 \end{bmatrix}$$



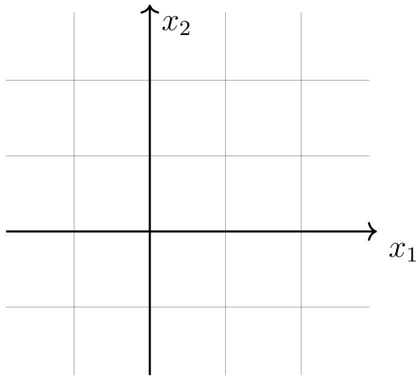
Definition 17. *Suppose that $B = b_1, \cdots b_p$ is a basis for a subspace H, **the coordinate of x relative to the basis** are the weights $c_1, \cdots c_p$ such that $x = c_1 b_1 + \cdots c_p b_p$, and the vector in $\mathbb{R}^p$*

$$[x]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

*is called the **coordinate of x (relative to B)** or the **B-coordinate vector of x.***

Exercise: Show that a B-coordinate vector of x is unique (assume two B-coordinates, and show that these two are necessarily equal).

Ho can we transfer B-coordinates to standard coordinates?



Example 10.4. *B-coordinates of the vector x be:*

$$[x]_B = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad B = \left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right)$$

Find the coordinates of x.

In general if B-coordinates of x is $[x]_B$ and $B = \{b_1, \cdots b_n\}$ is a basis, one can find the coordinates of x with respect to the standard bases (columns of I_n) with the following formula:

$$x = [b_1 \quad \cdots \quad b_n][x]_B$$

Inversely, B-coordinates of a given vector x can be written as:

$$[x]_B = [b_1 \quad \cdots \quad b_n]^{-1} x$$

11 Determinants

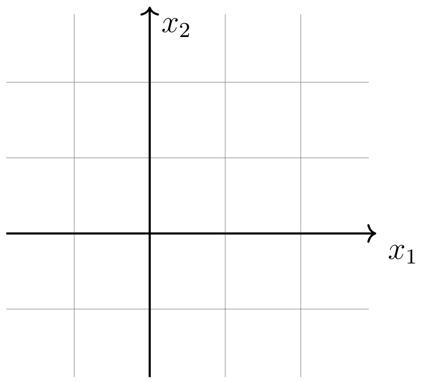
For the remainder of this course, we are going to work with square matrices $A \in \mathbb{R}^{n \times n}$. The determinant is a single number, that compresses a lot of information about an entire matrix. We will use it as another test for invertibility and singularity:

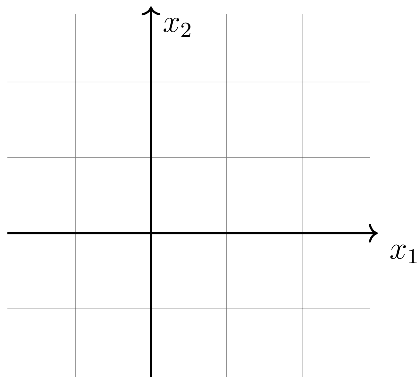
- $\det A = 0 \longrightarrow$

- $\det A \neq 0 \longrightarrow$

An interpretation of that is used in multivariable calculus is as follows:

If we apply A to the unit cube in $\mathbb{R}^n$, the det(A) gives the $n-$dimensional volume of the output.





Definition 18. *The function det: $\mathbb{R}^{n\times n} \mapsto \mathbb{R}$ with the properties:*

A.

B.

C.

*is called the **determinant** of a $n \times n$ matrix.*

Some important properties of the determinant of $A \in \mathbb{R}^{n\times n}$ are as follows:

(I) If A has a row of zeros then
proof.

(II) if A has two equal rows then
proof.

(III) If we add a multiple of one row to another row, then

proof.

(IV) If A is upper triangular, then

$$\det A = \left| \quad \right|$$

proof.

(V) $\det A = 0$ if and only if A is singular. $\det A \neq 0$ if and only if A is invertible.

proof.

(VI) The determinant of a 2×2 matrix is

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

proof.

(VII) For two matrices $A, B \in \mathbb{R}^{n \times n}$, $\det(AB) =$

(VIII) If A is invertible, then $\det(A^{-1}) =$

proof.

(IX) for any exponent $p \in \mathbb{N}, \det(A^p) =$

(X) For $\lambda \in \mathbb{R}, \det(\lambda A) =$

(XI) $\det(A^T) =$

Conclusion: In this introductory section, we have defined the determinant as a function that satisfies the three properties $(A), (B)$ and (C): the determinant of the identity matrix is 1, every row exchange reverses the sign of the determinant and the determinant is a linear map with respect to one fixed row. Of the properties derived from (A), (B) and (C), the most important ones are the invertibility criterion (V), the product formula (VII) and the fact (XI) that the transpose has no effect on the determinant.

11.1 How to compute the determinant in practice?

One of these three rules can be used:

1. formula for a 2×2 matrix.

2. forward elimination

3. co factor expansion

Theorem 10. *(Determinant by Gaussian elimination) Let $A \in \mathbb{R}^{n \times n}$. Then*

$$\textit{det}(A) =$$

Theorem 11. *(Derivation of the Cofactor Formula for the Determinant):*

Idea: Use property (C) to split the matrix into "basic matrices", which contain exactly on entry in every row and every column, all other entries being zero. Then, by row exchanges (using property (B)), every "basic matrix" can be transformed into a diagonal entries (by properties). Factoring out all the entries of one particular row or one particular column will then yield the cofactor formula.
2×2 *Case:*

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} & \\ c & d \end{vmatrix} + \begin{vmatrix} & \\ c & d \end{vmatrix}$$

$$= \begin{vmatrix} & \\ & \end{vmatrix} + \begin{vmatrix} & \\ & \end{vmatrix} + \begin{vmatrix} & \\ & \end{vmatrix} + \begin{vmatrix} & \\ & \end{vmatrix}$$

$$=$$

3×3 *Case:*

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} & & \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} & & \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} & & \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= \begin{vmatrix} & & \\ & & \\ & & \end{vmatrix} + \begin{vmatrix} & & \\ & & \\ & & \end{vmatrix} + \begin{vmatrix} & & \\ & & \\ & & \end{vmatrix}$$

$$+ \begin{vmatrix} & & \\ & & \\ & & \end{vmatrix} + \begin{vmatrix} & & \\ & & \\ & & \end{vmatrix} + \begin{vmatrix} & & \\ & & \\ & & \end{vmatrix}$$

$$=$$

$$= (\qquad\qquad)\, a_{11} + (\qquad\qquad)\, a_{12} + (\qquad\qquad)\, a_{13}$$

$$= \begin{vmatrix} +a_{11} & & \\ & & \\ & & \end{vmatrix} + \begin{vmatrix} & -a_{12} & \\ & & \\ & & \end{vmatrix} + \begin{vmatrix} & & +a_{13} \\ & & \\ & & \end{vmatrix}$$

$n \times n$ *Case:*

Theorem 12. *(Determinant by Cofactor Expansion)*

1. *Cofactor expansion across row i:*

$$\text{det } A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

2. *Cofactor expansion across column j:*

$$\text{det } A = a_{a_1 j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

If A_{ij} is the matrix that is formed by deleting row i and column j from A, then the cofactor is

$$C_{ij} = (-1)^{i+j} \det A_{ij}.$$

$$signs = \begin{vmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{vmatrix} \qquad A = \begin{vmatrix} \star & & \star & \star \\ \star & & \star & \star \\ & a_{ij} & & \\ \star & & \star & \star \end{vmatrix}$$

Example 11.1.

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 2 & 4 & 6 & 7 \\ -1 & -2 & 2 & 2 \end{vmatrix}$$

$$= \begin{vmatrix} & & & \\ & & & \\ & & & \\ & & & \end{vmatrix}$$

$$= \begin{vmatrix} & & & \\ & & & \\ & & & \\ & & & \end{vmatrix}$$

Example 11.2.

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 3 \\ 2 & 4 & 0 & 7 \\ -1 & -2 & 0 & 2 \end{vmatrix}$$

Conclusions: Our default method for calculating determinants is forward elimination. With every row exchange, the determinant changes sign. Only in two cases we use a different method:

- If the matrix is 2×2, then we use the formula: $ad - bc$.

- If the matrix is 3×3, or if it is larger but it has many rows and/or columns with many zeros, the cofactor expansion may be faster than elimination.

12 Eigenvalues and Eigenvectors

We already know one powerful set of numbers associated with a (rectangular) matrix:

-

In this chapter, we will introduce another set of numbers for a (square) matrix, which is even more relevant for a huge range of applications in mathematics, physics, computer science, engineering and economics:

-

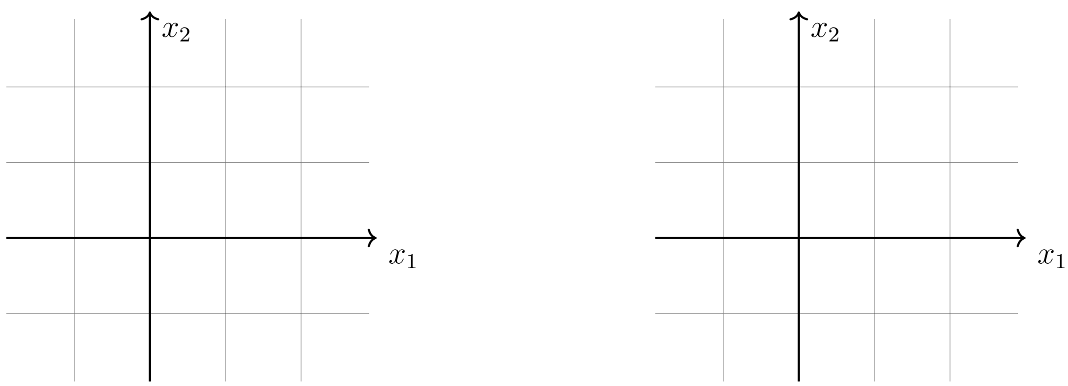
Definition 19. *(Eigenvalues and eigen vectors) Let* $A \in \mathbb{R}^{n \times n}$*. A nonzero vector* $v \in \mathbb{R}^n \backslash \{0\}$ *is said to be an eigenvector of* A *if:*

for some eigenvalue $\lambda \in \mathbb{R}$*.*

Example 12.1. *Check that* $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *is an eigenvector for matrix A ?*

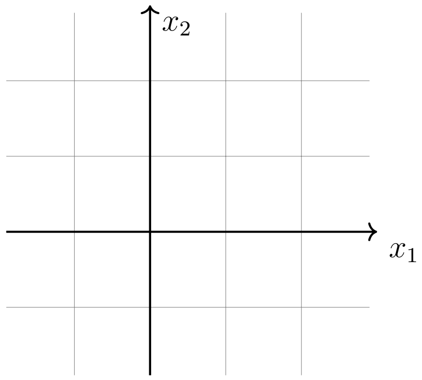
$$A = \begin{pmatrix} 0 & -2 \\ -4 & 2 \end{pmatrix}$$

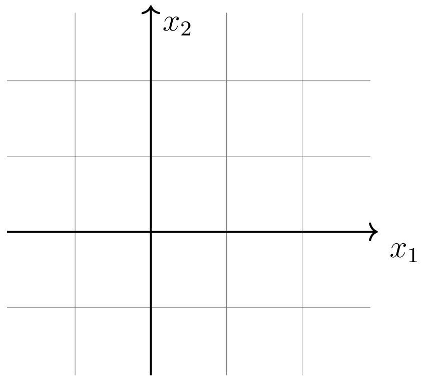
Example 12.2. *(Illustration of eigenvalues and eigenvectors) (a) projection onto a line*



Remark: Why fo we need Eigenvalues and Eigenvectors ? Remember about the linear transformations corresponding to contraction/dilation (Example 5.4)? Those are simple linear transformations which are represented by a diagonal matrix and can be easily understood by their geometrical representations:

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \qquad A : \mathbb{R}^2 \mapsto \mathbb{R}^2 \qquad A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 3x_2 \end{bmatrix}$$



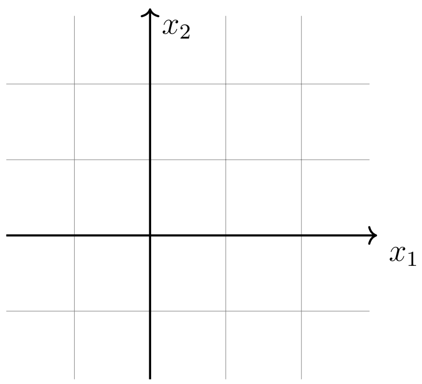


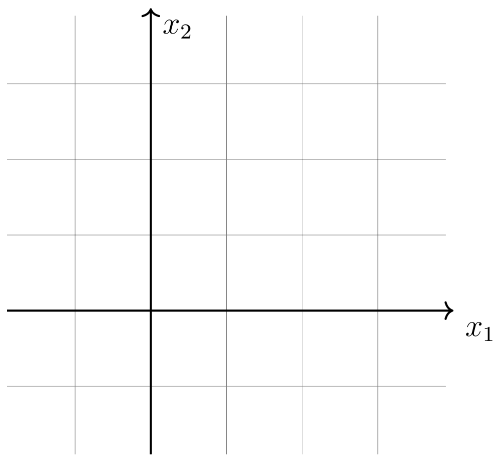
Now if we have a non-diagonal matrix A, eigenvectors and eigenvalues help us to better understand the "action" of A on the input vectors:

$$A = \begin{pmatrix} \frac{9}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{6}{5} \end{pmatrix} \qquad A : \mathbb{R}^2 \mapsto \mathbb{R}^2$$

Eigenvalues: $\lambda_1 = 2, \ \lambda_2 = 1$

Eigenvectors: $v_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$





How can we find the eigenvalues and eigenvectors ? Let's first start with the eigenvalues:

Example 12.3. *Find the eigenvalues of A.*

$$A = \begin{pmatrix} \frac{9}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{6}{5} \end{pmatrix}$$

Theorem 13. *A scalar $\lambda \in \mathbb{R}$ is an eigenvalue of a matrix $A \in \mathbb{R}^{n\times n}$, if and only if:*

$$\textit{det}(A - \lambda I_n) = 0$$

Example 12.4. *Find the eigenvalues of A.*

$$A = \begin{pmatrix} 5 & 0 & 3 \\ 1 & 2 & 1 \\ 3 & 0 & 5 \end{pmatrix}$$

Remark: The degree of the characteristic polynomial (i.e. $det(A - \lambda I_n)$) is equal to the size of A. Thus, the characteristic polynomial has at most n roots.
Sometimes it might have less than n roots and sometimes it might have no roots.

Some practical tips to find the roots of the characteristic polynomial:

1. If the degree is 2, use the quadratic formula:

$$a\lambda^2 + b\lambda + c = 0$$

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

2. If there is no constant, factor out as many λs as possible:

3. As in the example 12.4, one root might be explicitly visible:

4. Try if $\lambda = 1, \ \lambda = 2, ...$ is a root. Then we can factor out $(\lambda - p)$, p being the guessed root: $p(\lambda) = (\lambda - p)(...)$

5. If A is upper triangular, then the eigenvalues are the entries in the diagonal positions:

$$\det\,(A - I_n\lambda) = \begin{vmatrix} & & \\ & & \\ & & \end{vmatrix}$$

Attention: You are not allowed to do row reduction on A before computing the eigenvalues. Row reduction usually changes the eigenvalues of a matrix.

Definition 20. *Algebraic multiplicity is a property of the eigenvalue λ_i, and it denotes the multiplicity of λ_i as a root in the characteristic polynomial.*

Example 12.5. *Find the eigenvalues and their corresponding algebraic multiplicity for the following characteristic polynomial:*

$\lambda^4(\lambda - 2)(\lambda - 3) = 0$

Now let's turn our attention to the eigenvectors:

Example 12.6. *For the following matrix A, $\lambda_1 = 2$ is an eigenvalue. Find all of the eigenvectors corresponding to $\lambda_1 = 2$.*

$A = \begin{pmatrix} \frac{9}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{6}{5} \end{pmatrix}$

Definition 21. *The span of all eigenvectors with the same eigenvalues λ is called the eigenspace of A corresponding to λ, denoted by eig_λ. In other words, eig_λ is equal to the set of all solutions to $Ax = \lambda x$ (including 0)*

Theorem 14. *(Eigenspace and Nullspaces) If $\lambda \in \mathbb{R}$ is an eigenvalue of $A \in \mathbb{R}^{n \times n}$, then*

$$eig_\lambda A = nul(A - \lambda I_n)$$

Conclusions:

- If a matrix $A \in \mathbb{R}^{n \times n}$ is applied to any vector $x \in \mathbb{R}^n$, then the vector Ax usually has a different length than x and points in a different direction than x.
- If we can find a special vector x that does not change direction when we apply A to it, then this vector x is an eigenvector. The eigenvalue is the scaling factor λ that turns x into $Ax = \lambda x$
- If λ is an eigenvalue of the matrix A, then the matrix $A - \lambda I_n$ must be singular. Therefore:
  - its determinant is zero: $\det(A - \lambda I_n) = 0$
  - its nulspace nul($A - \lambda I_n$) = eig$_\lambda A$ is larger than just $\{0\}$

Conclusions (continued):

- Finding all eigenvalues and corresponding eigenvectors (eigenspaces) for the matrix A is a two-step procedure:

  1. Find all eigenvalues $\lambda_1, ..., \lambda_N$ by computing the roots of the characteristic polynomial:

  $$\det(A - \lambda I_n) = 0 \qquad\qquad \text{at most n real } \lambda\text{s}$$

  2. For each eigenvalue $\lambda_i, i \in \{1, \cdots, N\}$, find a basis E_i for the corresponding $\mathrm{eig}_{\lambda_i} A$:

  $$E_i = (v_i^1, v_i^2, \cdots, v_i^g) \qquad \text{at least one eigenvector}$$

  by computing all the "special solutions" of:

  $$(A - \lambda I_n)v_i = 0$$

- If A is a $n \times n$ matrix, then A is invertible if and only if the number 0 is **not** an eigenvalue.

Example 12.7. *Find all eigenvalues and eigenvectors of*

$$A = \begin{pmatrix} 2 & 3 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

- 1st step: Eigenvalues

- 2nd step: Eigenvectors

  - for $\lambda_{1,2} =$

- for $\lambda_3 =$

Definition 22. *(Geometric multiplicity of an Eigenvalue) Let $\lambda_i \in \mathbb{R}$ be an eigenvalue of a matrix $A \in \mathbb{R}^{n\times n}$, the geometric multiplicity g_i of λ_i is the dimension of its Eigenspace $\mathit{eig}_{\lambda_i} A$*

Theorem 15. *($1 \leq$ Geometric Multiplicity $\leq$ Algebraic Multiplicity) the geometric multiplicity g_i of the eigenvalues λ_i is at least 1 and at most equal to the algebraic multiplicity a_i:*

$$1 \leq g_i \leq a_i$$

Example 12.8. *Find the eigenvalues and eigenspaces, and specify the algebraic and geometric multiplicity of the following matrices. In addition, specify if the collection of eigenbases form an eigenvector basis for $\mathbb{R}^3$?*

(a) $A = \begin{pmatrix} 5 & 0 & 3 \\ 1 & 2 & 1 \\ 3 & 0 & 5 \end{pmatrix}$

(b) $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

Theorem 16. *(Linearly independent eigenvectors) If* $v_1, \cdots v_r \in \mathbb{R}^n$ *are eigenvectors corresponding to distinct eigenvalues* $\lambda_1, \cdots, \lambda_r \in \mathbb{R}$*, then these eigenvectors are linearly independent. (Proof on page 150 in textbook)*

Theorem 17. *(Eigenvectors that span $\mathbb{R}^m$) The following statements are equivalent:*

- *The characteristic polynomial of $A \in \mathbb{R}^{n \times n}$ can be decomposed into a product of linear factors:*

  $$det(A - \lambda I_n) = (\lambda_1 - \lambda)^{a_1}(\lambda_2 - \lambda)^{a_2} \cdots (\lambda_r - \lambda)^{a_r}$$

  and all r distinct eigenvalues $\lambda_1, \cdots, \lambda_r \in \mathbb{R}$ have

  $$a_i = g_i \qquad\qquad (i = 1, \cdots, r)$$

  $$\textit{algebraic multiplicity} = \textit{geometric multiplicity}$$

- *The n eigenvectors $v_1^1, \cdots, v_1^{g_1}, v_2^1, \cdots, v_2^{g_2}, \cdots, v_r^1, \cdots v_r^{g_r}$ span the full space $\mathbb{R}^n$.*

- *These eigenvectors form a basis for $\mathbb{R}^n$, a so-called eigenvector basis.*

12.1 Diagonalization

Diagonal matrices are very easy to deal with: The inverse of a diagonal matrix is simply:

(Other) powers of diagonal matrices are also straightforward to evaluate:

We will now use eigenvalues and eigenvectors to transform a matrix $A \in R^{n \times n}$ to a diagonal matrix $D \in \mathbb{R}^{n \times n}$, if possible, which will allow for very simple calculations.

Definition 23. *(Similar matrices): Two matrices A and B are said to be similar, if there exists an invertible matrix $P \in \mathbb{R}^{n \times n}$ such that:*

$$A = PBP^{-1}$$

Theorem 18. *If A and B are similar, then A and B have the same characteristic polynomial and hence the same eigenvalues.*

Definition 24. *A matrix is said to be diagonalisable, if it is similar to a diagonal matrix: there exist an invertible matrix $P \in \mathbb{R}^{n \times n}$ and a diagonal matrix $D \in \mathbb{R}^{n \times n}$ such that:*

$$A = PDP^{-1}$$

Example 12.9. *(Power of a matrix) We define:*

$$P = \begin{pmatrix} 2 & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{pmatrix}, \qquad D = \begin{pmatrix} -1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \qquad A = PDP^{-1} = \begin{pmatrix} -\frac{5}{2} & -3 \\ \frac{3}{2} & 2 \end{pmatrix}$$

Derive a formula for A^k, where $k \in \mathbb{N}$ is any positive integer.

Theorem 19. *(Continuation of Theorem 17) The following statements are equivalent:*

- *The characteristic polynomial of $A \in \mathbb{R}^{n\times n}$ can be decomposed into a product of linear factors:*

  $$det(A - \lambda I_n) = (\lambda_1 - \lambda)^{a_1}(\lambda_2 - \lambda)^{a_2} \cdots (\lambda_r - \lambda)^{a_r}$$

  and all r distinct eigenvalues $\lambda_1, \cdots, \lambda_r \in \mathbb{R}$ have

  $$a_i = g_i \qquad\qquad (i = 1, \cdots, r)$$

  algebraic multiplicity = geometric multiplicity

- *The matrix A is diagonalisable.*
  In this case we have $A = PDP^{-1}$ with

Example 12.10. *(Denationalization) Determine whether or not the matrix*

$$(a) \quad A = \begin{pmatrix} \frac{4}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{pmatrix} \qquad\qquad (b) \quad A = \begin{pmatrix} 2 & 3 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

is diagonalisable. If so, write down a corresponding diagonal matrix D and a transformation matrix P such that $A = PDP^{-1}$.

Conclusions: Denationalisation allows us to extend the very simple calculations with diagonal matrices to diagonalisable matrices:

- $A \in \mathbb{R}^{n\times n}$ is diagonalisable if and only if it has n real eigenvalues (not necessarily distinct) and n linearly independent eigenvectors, i.e. an eigenvector basis for $\mathbb{R}^n$.
- In the similarity transformation $A = PDP^{-1}, D \in \mathbb{R}^{n\times n}$ is a matrix with the eigenvalues of A on the diagonal, $P \in \mathbb{R}^{n\times n}$ has the corresponding eigenvectors as columns.
- Powers of a diagonalisable matrix A can easily be calculated as $A^k = PD^kP^{-1}$.

Example 12.11. *(Predetor-prey system) A dynamical system is defined by the following equations: (please refer to the dynamical systems lecture notes, page 27, for a complete description of the problem)*

$$g_{n+1} = 0.38g_n + 0.24y_n$$
$$y_{n+1} = -0.36g_n + 1.22y_n$$

Or in matrix form:

$$f_{n+1} = \begin{pmatrix} 0.38 & 0.24 \\ -0.36 & 1.22 \end{pmatrix} f_n, \qquad f_0 = \begin{bmatrix} g_0 \\ y_0 \end{bmatrix}$$

Find the general formula for f_{n+1} in terms of f_0, with the method of diagonalization:

Summery and conclusions of Discrete Dynamical Systems: The linear discrete dynamical system is described by the formula $x_{n+1} = Ax_n$ with an initial condition x_0 and $A \in \mathbb{R}^{k\times k}$. With the help of eigenvalues/spaces/vectors, we can achieve the following two goals:

(a) Find an explicit formula for x_{n+1} in terms of n and x_0. Depending on the problem, x_0 might be given as a vector of numbers or simply given in parametric form. Remind that the explicit formula should not include x_n (that is why it is called explicit after all).

(b) Find the general behavior of the discrete dynamical system in long-run, based on different values of x_0.

In doing so the following steps should be taken:

Step 1: Find all the eigenvalues and eigenspaces of A.

Step 2: Check for each eigenvlaue λ_i, if the geometric multiplicity and algebraic multiplicity are equal ($a_i = g_i$). If this condition does not hold at least for one eigenvalue, the method breaks down. It means that we cannot find an explicit formula for f_{n+1}, for any initial condition x_0. If this condition holds for all eigenvalues, we can find the eigenvector basis which spans the full space, i.e. a set which contains all bases corresponding to all eigenspaces $P = (v_1, \cdots, v_k)$ (refer to the Theorem 17 to recall about the eigenvector basis)

Step 3: Now we need to decide which method we prefer to choose in order to find the general solution:

(a) *Method of undetermined coefficients:* If the initial condition is not given in terms of numbers, there is not much left to do. The explicit solution can be written in the following form (to see why, refer to page 30 in the discrete dynamical system lecture note):

$$x_n = c_1\lambda_1^n v_1 + \cdots + c_k\lambda_1^n v_k \tag{26}$$

In the above equation, only $c_1, \cdots, c_k$ are unknown (recall that n is only the time variable, and the general solution will always depend on it). They will remain unknown if we do not know about the initial condition. If x_0 is given, then it can be expressed as a linear combination of the eigenvector basis (Why?):

$$x_0 = c_1v_1 + c_2v_2 + \cdots + c_kv_k$$

The above vector equation is simply a linear system with k unknown variables $c_1, \cdots, c_k$. After finding those unknown variables, the general solution 26 is explicitly expressed in terms of n.

(b) *Method of diagonalization* In this method, we first diagonalize the matrix A (refer to the Theorem 19). Since we already know the eigenvector basis, we can write:

$$A = PDP^{-1},$$

Next, the general solution can be directly expressed as:

$$x_{n+1} = PD^nP^{-1}x_0$$

. In the above vector equation, the only unknown variable is n, which is the time variable).

Step 4: Describing the general behavior of the system in long run: First, we need to plot the phase diagram (a diagram whose axes are each component of our input vector) with all eigenspaces. We want to know if we start from any initial vector, what will be the behavior of the system in the future. Starting from an initial vector on the eigenspace, x_n either gets attracted by the origin (if $|\lambda| < 1$), or it gets repelled by the origin (if $|\lambda| > 1$). If the initial vector is not on the eigenspace, we look at the explicit general formula that we found on the previous step in order to predict the behavior of the system. In short, any arbitrary initial vector gets attracted by the eigenspaces with $|\lambda| > 1$ and get repelled by the eigenspaces with $|\lambda| < 1$, and are neutral with respect to eigenspaces with $|\lambda| = 1$.

13 Orthogonality and Least Squares

Vectors allow us to use analytical tools to solve geometric problems and to extend the notions of length and angles to very general vector spaces. All the necessary information is contained in the

Definition 25. *(Dot product on $\mathbb{R}^n$) The dot product of two vectors $x, y \in \mathbb{R}^n$ is defined by*

$$x.y = x_1y_1 + x_2y_2 + \cdots + x_ny_n \in \mathbb{R}$$

Theorem 20. *(Properties of the dot product) The dot product on $\mathbb{R}^n$ is*

(a) positive definite:

(b) symmetric:

(c) linear in each argument:

Definition 26. *(Euclidean Norm on $\mathbb{R}^n$) The (Euclidean norm), 2-norm or length of a vector is defined by:*

Theorem 21. *(Properties of the Euclidean Norm) The Euclidean norm on $\mathbb{R}^n$ is*

(a) positive definite

(b) absolutely homogeneous

(c) subadditive

A vector is called a unit vector if $||x|| = 1$.

Example 13.1. *(Normalization) We define* $x = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$ *and* $\hat{x} =$

$$||x|| =$$

$$||\hat{x}|| =$$

Example 13.2. *(Distance between Two Vectors) Find the distance between* $x = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$ *and* $y = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$*.*

Theorem 22. *For two vectors* $x, y \in \mathbb{R}^n$*, that span the angle* θ*:*

Theorem 23. *(Angle between two vectors) Find the angle between* $x = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$ *and* $y = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$*.*

Definition 27. *(Orthogonality) Two vectors* $x, y \in \mathbb{R}^n$ *are said to be orthogonal if*

Definition 28. *(Orthogonal Complement) Let H be a subspace of $\mathbb{R}^n$. The set of all vectors that are orthogonal to all vectors in H:*

is called the orthogonal complement of H.

Example 13.3. *(Orthogonal complement of a plane)*

- *in $\mathbb{R}^2$: Let $H = span\left(\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\2\end{pmatrix}\right) =$*

- *in $\mathbb{R}^3$: Let $H = span\left(\begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}3\\2\\1\end{pmatrix}\right)$*

Theorem 24. *(Properties of Orthogonal Complement)*

(a) The orthogonal complement $H^\perp$ of any subspace is a subspace as well.

(b) If $H \in \mathbb{R}^n$ is a subspace, then

$$(H^\perp)^\perp = H$$

(c) Let $A \in \mathbb{R}^{m\times n}$:

$$(Row\ A)^\perp = Nul\ A$$

$$(Col\ A)^\perp = Nul\ (A^T)$$

13.1 What is the significance of Orthogonal and Orthonormal bases?

In this section, we're revisiting the problems of finding the coordinates of a vector $x \in H$ in terms of a given basis $B = (u_1, \cdots u_r)$ of the space $H \subseteq \mathbb{R}^n$: find weights (B-coordinates) $c_1, \cdots, c_r \in \mathbb{R}$ such that

$$x = c_1 u_1 + \cdots + c_r u_r$$

the solution is given by the solution of the linear system:

For orthogonal and orthonormal bases, the solution is a lot shorter.

Definition 29. *(Orthogonal and orthonormal sets and bases) A set $\{u_1, \cdots, u_r\} \subseteq \mathbb{R}^n$ or a basis $(u_1, \cdots, u_r) \subseteq \mathbb{R}^n$ of a subspace H is said to be*

- *orthogonal, if*

- *orthonormal, if*

Theorem 25. *(Nonzero orthogonal sets are linearly independent) if $\{u_1, \cdots, u_r\} \subseteq \mathbb{R}^n \backslash \{0\}$ is an orthogonal set, then the vectors $u_1, \cdots, u_r$ must be linearly independent.*
Proof.

Example 13.4. *(Orthogonal basis for* $\mathbb{R}^3$ *) Show that*

$$\mathcal{U} = \left(\begin{pmatrix} 0 \\ 4 \\ 3 \end{pmatrix}, \begin{pmatrix} -20 \\ -9 \\ 12 \end{pmatrix}, \begin{pmatrix} -15 \\ 12 \\ -16 \end{pmatrix} \right)$$

is an orthogonal basis for $\mathbb{R}^3$*.*

Example 13.5. *(Orthonormal bases for* $\mathbb{R}^3$*) Examples for orthonormal bases of* $\mathbb{R}^3$ *are:*

Theorem 26. *(Coordinates in orthogonal and orthonormal bases) If* $\mathcal{U} = (u_1, \cdots, u_r) \subseteq \mathbb{R}^n$ *is an orthogonal basis of a subspace* H*, then any vector* $x \in H$ *has a unique representation in that basis:*

If $\mathcal{U}$ *is even orthonormal, then*

proof.

Example 13.6. *(Coordinates in an orthogonal basis) In the standard basis* $\mathcal{E} = (e_1, e_2, e_3)$ *of* $\mathbb{R}^3$*, a vector is given as*

$$x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Find its coordinates in the basis $\mathcal{U}$ *from example 13.4*

Theorem 27. *(Matrix with orthonormal columns) a matrix* $U \in \mathbb{R}^{m \times n}$ *has orthonormal columns if and only if* $U^T U = I_n$*.*
Proof.

Definition 30. *(Orthogonal matrix) A square matrix* $U \in \mathbb{R}^{n \times n}$ *is called orthogonal if*

Theorem 28. *(Angle and Length-Preserving Linear Transformation) If $U \in \mathbb{R}^{m\times n}$ has orthonormal columns, then*

13.2 How to project a vector orthogonally onto a subspace

Given some vector $x \in \mathbb{R}^n$, how can we find its projection onto a subspace $H \subseteq \mathbb{R}^n$, e.g. onto a line/plane/hyperspace through the origin ?

Theorem 29. *(Orthogonal decomposition) Let $H \subset \mathbb{R}^n$ be a subspace.*

- *Each vector $x \in \mathbb{R}^n$ can be decomposed into the sum of a vector $\hat{x} \in H$ and a vector $e \in H^\perp$*

  $$x = \hat{x} + e$$

  This orthogonal decomposition is unique.

- *If $(u_1, \cdots, u_r)$ is an orthogonal basis for H, then*

  $$\hat{x} =$$

  $$e =$$

Theorem 30. *Let $H \subseteq \mathbb{R}^n$ be a subspace, $x \in \mathbb{R}^n$ and $\hat{x} \in H$ the projection of x onto H. Then $\hat{x}$ is the closest point in H to x:*

$$\forall h \in H: \qquad ||x - \hat{x}|| \leq ||x - h||$$

Example 13.7. *(Calculating the Orthogonal Projection) Let H be the plane spanned by the two orthogonal vectors*

$$u_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \qquad \text{and} \qquad u_2 = \begin{pmatrix} -5 \\ 2 \\ 1 \end{pmatrix}$$

Find the point $\hat{x}$ on the plane H which is closest to the point

$$x = \begin{pmatrix} -7 \\ 9 \\ 7 \end{pmatrix}$$

and calculate the distance of x from H.

Example 13.8. *Consider the plane $x - y + z = 0$.*

(a) Find the 3×3 matrix T_1 which represents projection of $\mathbb{R}^3$ onto a vector orthogonal to this plane.
Solution:
In general, an orthogonal vector to any plane in $\mathbb{R}^3$ written in the form $ax + by + cz = d$is given as: $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$. Thus, in this particular example, the orthogonal vector is:

$$u_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

We are asked to find a linear transformation $T_2(x) : \mathbb{R}^3 \mapsto \mathbb{R}^3$ which projects any point on the subspace spanned by u_1, that is a line in $\mathbb{R}^3$. Let's call this line the subspace H, and a basis for it is $\mathcal{U} = \{u_1\}$. Obviously, this is an orthogonal basis (Why?). As we learned previously about finding the standard matrix of a linear transformation, we need to find the "action" of the linear map on the columns of identity matrix. Thanks to the orthogonality of this basis, from Theorem 29 in the lecture notes (or Theorem 8 in the textbook) we have:

$$T_1(e_1) = \textit{proj}_H e_1 = \frac{u_1 \cdot e_1}{u_1 \cdot u_1} u_1 = \frac{1}{3}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$$T_1(e_2) = \textit{proj}_H e_2 = \frac{u_1 \cdot e_2}{u_1 \cdot u_1} u_1 = \frac{-1}{3}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$$T_1(e_3) = \textit{proj}_H e_3 = \frac{u_1 \cdot e_3}{u_1 \cdot u_1} u_1 = \frac{1}{3}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$$A_1 = [T_1(e_1) \qquad T_1(e_2) \qquad T_1(e_3)] = \frac{1}{3}\begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$$

(b) Let $a = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Find vectors v and w such that $a = v + w$, where v is in the plane and w is perpendicular to the plane.
Solution:

$$w = \textit{proj}_H a = \frac{a \cdot u_1}{u_1 \cdot u_1} u_1 = \frac{1}{3}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$$v = a - w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{4}{3} \\ \frac{2}{3} \end{pmatrix}$$

(c) Find the 3×3 matrix T_2 which represents projection of $\mathbb{R}^3$ onto this plane.
Solution:
Similar to (a), we want to find a the "action" of the linear map to the columns of

the identity matrix. As we did for vector a in part (b), we can find the projection of e_1, e_2, e_3 on the plane $x - y + z = 0$ by:

$$T_2(e_1) = e_1 - T_1(e_1)$$
$$T_2(e_2) = e_2 - T_1(e_2)$$
$$T_2(e_3) = e_3 - T_1(e_3)$$

$$A_2 = [T_2(e_1) \qquad T_2(e_2) \qquad T_2(e_3)] = I_3 - A_1 = \frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$$

13.3 How to find Approximate Solutions to Inconsistent Linear System?

With this b, the problem $Ax = b$ has no solution, since $b \notin \text{col}(A)$, i.e. it is impossible to make $||b - Ax|| = 0$. Instead, we will try to make $||b - Ax||$ as small as possible. To find such *approximate* solutions of $Ax = b$, we proceed as follows:

- Approximate the right hand side b with $\hat{b} = \text{proj}_{\text{col}(A)} b$

- Solve $Ax = \hat{b}$ instead. If the approximation error $||b - \hat{b}|| = ||b - Ax||$ is small, then $Ax \approx b$.

Definition 31. *(Least square solution) Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. A vector $\hat{x} \in \mathbb{R}^n$ is called a least-square solution of $Ax = b$ if*

$$\forall x \in \mathbb{R}^n : ||b - A\hat{x}|| \leq ||b - Ax||$$

How to find that approximate solution $\hat{x}$?

Example 13.9. *Let*

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 2 & 2 \end{pmatrix} \qquad\qquad b = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Check if the system $Ax = b$ is consistent. If not, find the approximate solution and the error.

Theorem 31. *The set of least square solutions of $Ax = b$ concides with nonempty set of solutions of the normal equations $A^T Ax = A^T b$.*

Remark: Equation $A^T Ax = A^T b$ always has a solution. However, this solution is not always unique:

Theorem 32. *The following statements are equivalent:*

- *The problem $Ax = b$ has a unique least-squares solution.*
- *The problem $A^T Ax = A^T b$ has a unique solution.*
- *nul*$(A) = 0$
- *A has full column rank n.*
- *The columns of A are linearly independent.*
- *The matrix $A^T A$ is invertible.*

Example 13.10. *Find the least-square solution of the linear system:*

$$\begin{pmatrix} 2 & 1 \\ -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \\ 1 \end{pmatrix}$$

Is this least-square solution unique?

Example 13.11. *(Calculating the Orthogonal Projection) Let H be the plane spanned by the two (non-orthogonal) vectors:*

$$a_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \qquad \textit{and} \qquad a_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

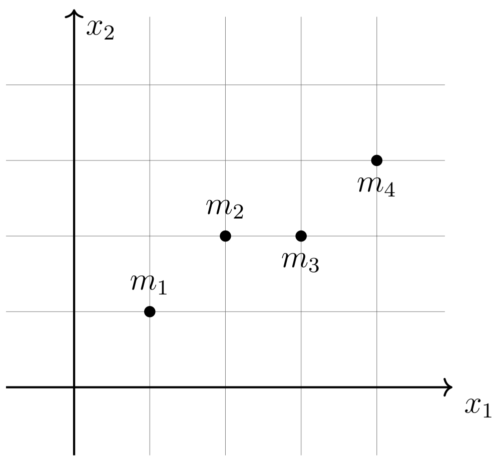
. Find the point $\hat{b}$ on the plane H which is closest to the point

$$b = \begin{pmatrix} 3 \\ 5 \\ 1 \end{pmatrix}$$

and calculate the distance of b from H.

Example 13.12. *(Least-square fitting) An experimental study has produced the following data:*

$$m_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad m_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$
$$m_3 = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \qquad m_4 = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$



Find the best linear fit with the least square error to these data.