MATH 33A – WINTER 2022

Pablo S. Ocal

based on "Linear Algebra with Applications" by Otto Bretscher.

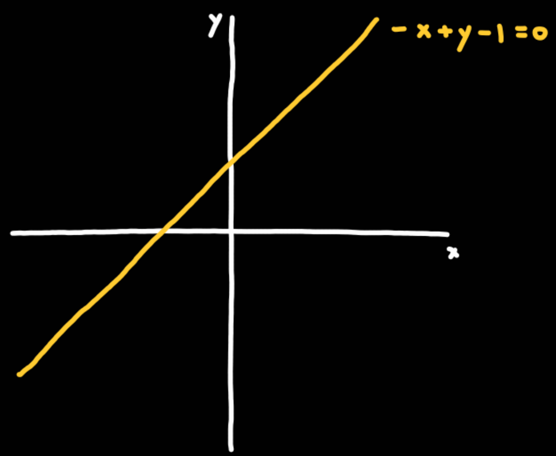
1. Introduction to Linear Algebra. (Chapter 1 and Chapter 2)

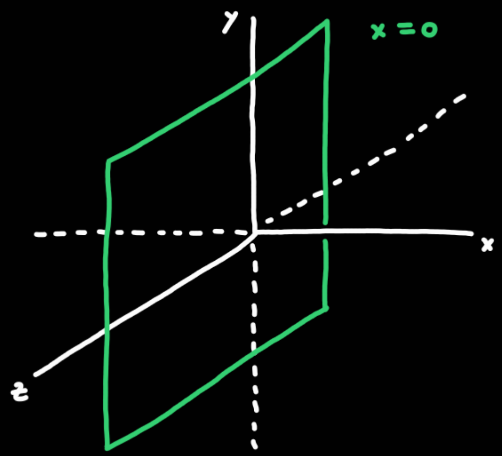
Linear algebra is the study of linear equations and linear transformations.

A linear equation has the form:

$$a_1 x_1 + a_2 x_2 + \dots + a_n x_n + b = 0,$$

where $a_1, \dots, a_n$ are real numbers called coefficients, $x_1, \dots, x_n$ are variables, and b is a real number called the constant term. Geometrically, linear equations define lines, planes, and objects that we will call subspaces.



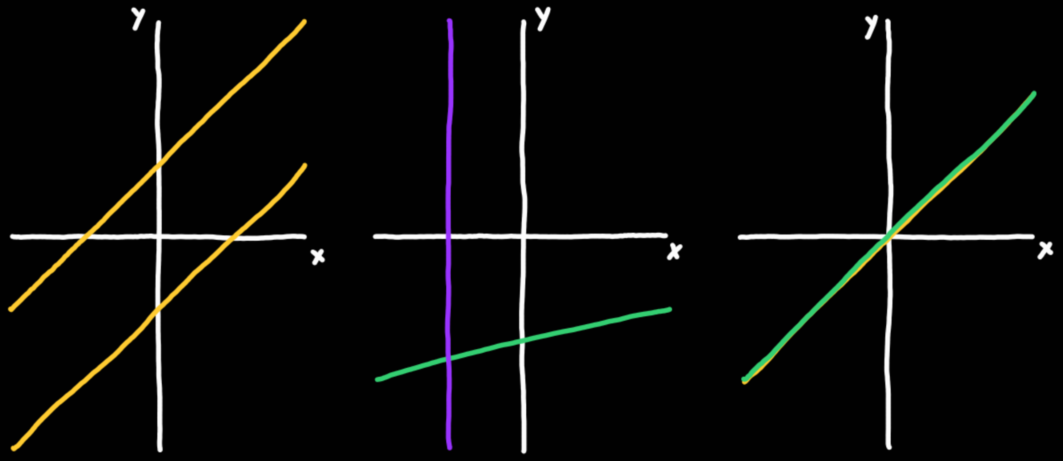


Systems of linear equations can have no solution, one solution, or infinitely many solutions.

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n + b_1 &= 0 \\ &\vdots \\ a_{m1} x_1 + \dots + a_{mn} x_n + b_m &= 0 \end{aligned}$$

$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n + b_n = 0$

This happens because the solutions are the intersection points of the geometric objects defined by the equations. There are either no intersections, one intersection, or infinitely many intersections.



parallel lines　　one intersection　　double line

To handle and solve systems of linear equations, we use <u>matrices</u>:

$$\begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix}$$ is an $n \times m$ matrix with entries a_{ij}.

A <u>matrix</u> is a rectangular array of numbers. If a matrix has n rows and m columns, we say that the <u>size</u> of the matrix is $n \times m$. We say that two matrices A and B are equal when their <u>entries</u> a_{ij} and b_{ij} are equal.

Some families of matrices receive special names:

(i) Square matrices.

(ii) Diagonal matrix.

(iii) Upper triangular matrix.

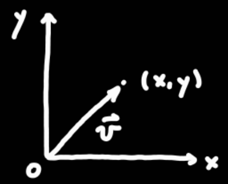
(iv) Lower diagonal matrix.

(v) Zero matrix.

A <u>vector</u> is a matrix with only one column. The entries of a vector are called its <u>components</u>.

The set of all column vectors with n components is denoted by $\mathbb{R}^n$. We will refer to $\mathbb{R}^n$ as a <u>vector space</u>.

The standard representation of a vector in the Cartesian coordinate plane is as an arrow from the origin: $\vec{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ is represented as



. However, thinking about vectors conceptually as a list of numbers written in a column will be useful.

Given a system of n linear equations in m variables:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m &= b_1 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m &= b_n
\end{aligned}
$$

we store the information on an <u>augmented matrix</u>:

$$
\left[\begin{array}{cccc|c}
a_{11} & a_{12} & \cdots & a_{1m} & b_1 \\
\vdots & \vdots & & \vdots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nm} & b_n
\end{array}\right]
$$

and simplify it using three <u>row operations</u>: (we will soon see that these correspond to multiplication by invertible matrices, specifically diagonal matrices and permutation matrices).

(1) Divide a row by a non-zero scalar.

(2) Subtract a multiple of a row from another row.

(3) Swap two rows.

Example:

The system of linear equations:

$$\begin{aligned} 2x + 8y + 4z &= 2 \\ 2x + 5y + z &= 5 \\ 4x + 10y - z &= 1 \end{aligned}$$

has augmented matrix

$$\left[\begin{array}{ccc|c} 2 & 8 & 4 & 2 \\ 2 & 5 & 1 & 5 \\ 4 & 10 & -1 & 1 \end{array}\right]$$

which can be simplified into:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 11 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

giving the solution $x = 11$, $y = -4$, $z = 3$.

The simplified form is called <u>reduced row-echelon form</u>, and solves the system of linear equations.

A matrix is in reduced row-echelon form if it satisfies all the following conditions:

(i) If a row has non-zero entries, then the first non-zero entry is a 1.

This is called the <u>leading</u> 1, or <u>pivot</u>, of the row.

(ii) If a column contains a leading 1, then all the other entries in the column are 0.

(iii) If a row contains a leading 1, then each row above it contains a leading 1 further to the left.

If there are rows of zeroes, by (iii), they must appear at the bottom of the matrix.

<u>Example:</u> The zero matrix is in reduced row-echelon form.

<u>Example:</u> When reducing the augmented matrices of three systems we obtain:

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

How many solutions are there in each case?

(a) No solutions. (b) Infinitely many solutions. (c) One solution.

A system of equations is called <u>consistent</u> if there is at least one solution, and <u>inconsistent</u> if there are no solutions.

<u>Theorem:</u> A linear system is inconsistent if and only if the reduced row-echelon form of its augmented matrix contains the row $[0 \cdots 0 1 \mid 1]$. If a linear system is consistent then:

(i) it has infinitely many solutions if there is at least one free variable, or

(ii) it has exactly one solution if all the variables are leading.

More useful information can be obtained from the reduced form of a matrix, like the rank.

The <u>rank</u> of a matrix A, denoted $\mathrm{rank}(A)$, is the number of leading 1's in $\mathrm{rref}(A)$.

<u>Example:</u> For $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ we have $\mathrm{rref}(A) = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ so $\mathrm{rank}(A) = 2$.

<u>Theorem:</u> Consider a system of n equations in m variables (so its coefficient matrix has size $n \times m$). Then: (why? justify this!)

(1) We have $\mathrm{rank}(A) \leq n$ and $\mathrm{rank}(A) \leq m$.

(2) If $\mathrm{rank}(A) = n$, then the system is consistent.

(3) If $\mathrm{rank}(A) = m$, then the system has at most one solution.

(4) If $\mathrm{rank}(A) < m$, then the system has either zero or infinitely many solutions.

<u>Example:</u>

1. Suppose we have a system with fewer equations than variables. How many solutions could we have? Answer: no solutions or infinitely many, since $\mathrm{rank}(A) \leq n < m$.

   That is, if a linear system has a unique solution, then there must be at least as many equations as variables.

2. Suppose we have a system with n equations and n variables. When do we have exactly one solution? Answer: if and only if the rank of the matrix is n.

Since matrices play such a big role in linear algebra, we have to get comfortable manipulating them. This includes addition of matrices, scalar multiples of matrices, and later multiplications.

<u>Addition:</u> The matrix $C = A + B$ has entries $c_{ij} = a_{ij} + b_{ij}$.

<u>Scalar multiplication:</u> The matrix $C = kA$ has entries $c = k a_{ij}$.

<u>Dot product:</u> The dot product of two vectors is a scalar: $\vec{x} \cdot \vec{y} = \sum_{i=1}^{n} x_i y_i$. (this is the precursor of matrix multiplication)

<u>Product of a matrix with a vector:</u>

$$A\vec{x} = \begin{bmatrix} - & \vec{w}_1 & - \\ & \vdots & \\ - & \vec{w}_n & - \end{bmatrix} \vec{x} = \begin{bmatrix} \vec{w}_1 \cdot \vec{x} \\ \vdots \\ \vec{w}_n \cdot \vec{x} \end{bmatrix}$$

$$A\vec{x} = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = x_1 \vec{v}_1 + \cdots + x_m \vec{v}_m$$

A has size $n \times m$, $\vec{x}, \vec{w}_1, \dots, \vec{w}_n$ are vectors in $\mathbb{R}^m$, $\vec{v}_1, \dots, \vec{v}_m$ are vectors in $\mathbb{R}^n$.

<u>Algebraic rules:</u> $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$ and $A(k\vec{x}) = kA\vec{x}$.

A vector $\vec{v}$ is a <u>linear combination</u> of the vectors $\vec{v}_1, \dots, \vec{v}_m$ in $\mathbb{R}^n$ if there are scalars $a_1, \dots, a_m$ such that $\vec{v} = a_1 \vec{v}_1 + \cdots + a_m \vec{v}_m$.

Given a linear system with augmented matrix $[A \mid \vec{b}]$, we can write it as an equality of matrices: $A\vec{x} = \vec{b}$ where $\vec{x}$ is the vector of variables.

<u>Example:</u>

The system of linear equations:

$$\begin{aligned} 2x + 8y + 4z &= 2 \\ 2x + 5y + z &= 5 \\ 4x + 10y - z &= 1 \end{aligned}$$

is equivalent to the equation

$$\begin{bmatrix} 2 & 8 & 4 \\ 2 & 5 & 1 \\ 4 & 10 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}$$

Example: Are the following statements true or false?

1. There exists a 3×4 matrix of rank 4. False!

2. There exists a system of three linear equations with three unknowns that has exactly three solutions. False!

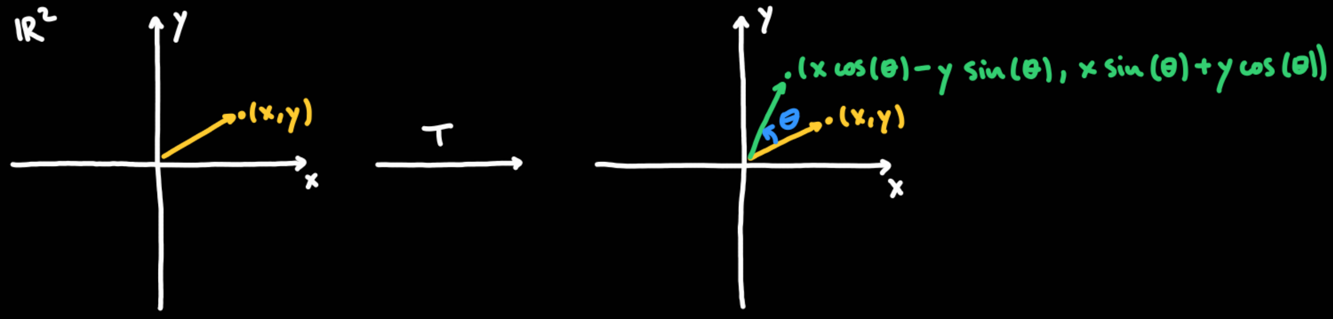
3. $\operatorname{rank}\begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} = 2.$ False!

4. If A is a 3×4 matrix of rank 3, then the system $A\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ must have infinitely many solutions. True!

A function T from X to Y is an assignment of an unique element y of Y to each element x of X. We call X the domain of T and Y the range of T.

A linear transformation is a function T from $\mathbb{R}^m$ to $\mathbb{R}^n$ such that there exists an $n \times m$ matrix A with $T(\vec{x}) = A\vec{x}$ for all $\vec{x}$ in $\mathbb{R}^m$.

Example: Consider the function from $\mathbb{R}^2$ to $\mathbb{R}^2$ given by a rotation of angle θ.

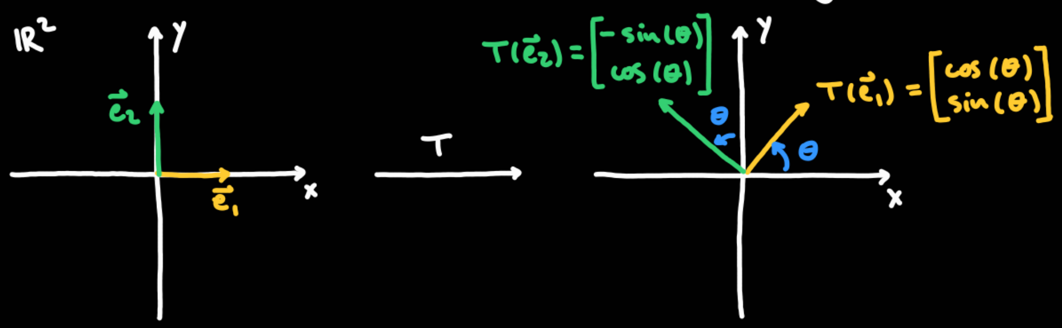


This rotation is a linear transformation because:

$$T(\vec{x}) = T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos(\theta) - y\sin(\theta) \\ x\sin(\theta) + y\cos(\theta) \end{bmatrix}.$$

__Theorem:__ Let T be a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$. The columns of the matrix associated to T are $T(\vec{e}_1), \dots, T(\vec{e}_m)$ where $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \vec{e}_m = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$.

__Example:__ Consider the function from $\mathbb{R}^2$ to $\mathbb{R}^2$ given by a rotation of angle θ.



So the matrix associated to T is $\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$.

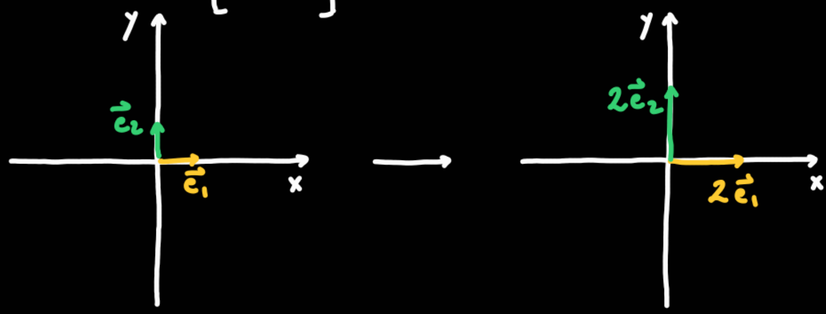
__Theorem:__ A function T from $\mathbb{R}^m$ to $\mathbb{R}^n$ is a linear transformation if and only if:

(i) $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$ for all $\vec{v}, \vec{w}$ in $\mathbb{R}^m$, and

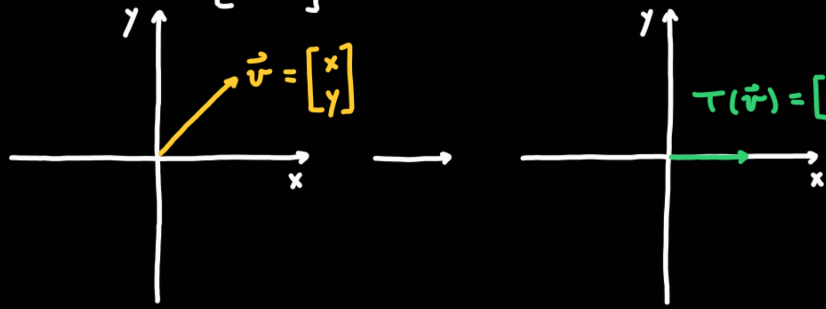
(ii) $T(\lambda \vec{v}) = \lambda T(\vec{v})$ for all $\vec{v}$ in $\mathbb{R}^m$ and λ in $\mathbb{R}$.

__Example:__

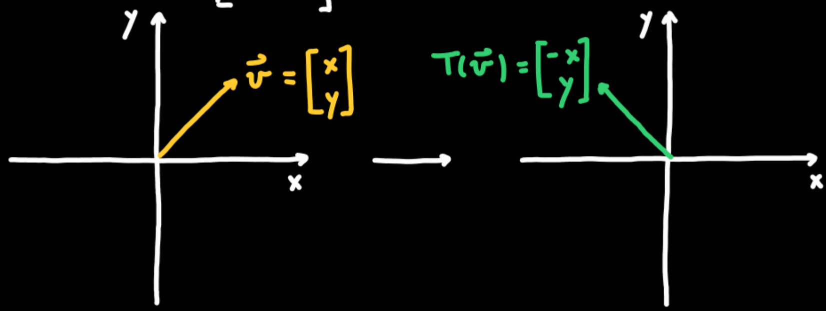
1. The matrix $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ is a dilation by 2 (or scaling).



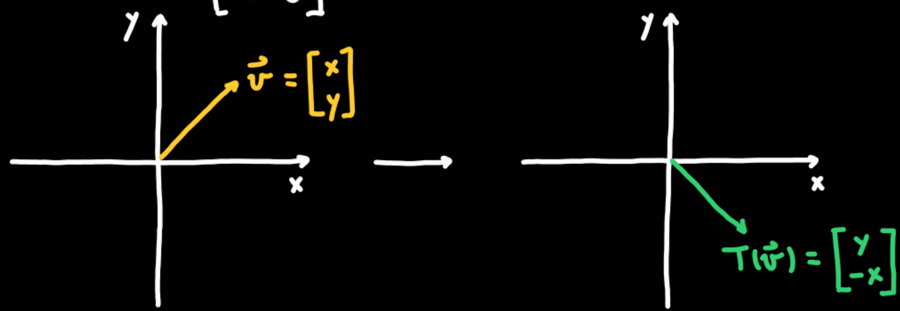
2. The matrix $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is an orthogonal projection onto the horizontal axis:



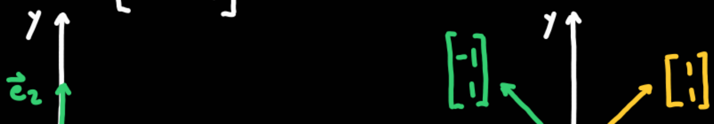
3. The matrix $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ is a reflection about the vertical axis:

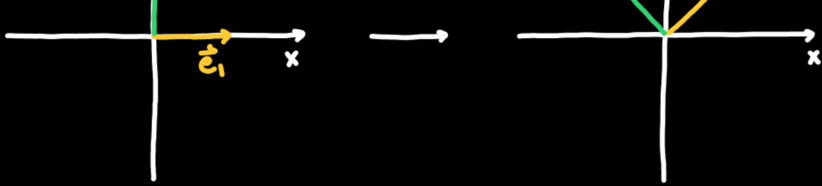


4. The matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is a clockwise rotation of $\frac{\pi}{2}$.



5. The matrix $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ is a rotation of $\frac{\pi}{4}$ and a dilation of $\sqrt{2}$.





Scaling.

Is given by multiplying a vector $\vec{x}$ with a diagonal matrix $\begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$, k in $\mathbb{R}$.

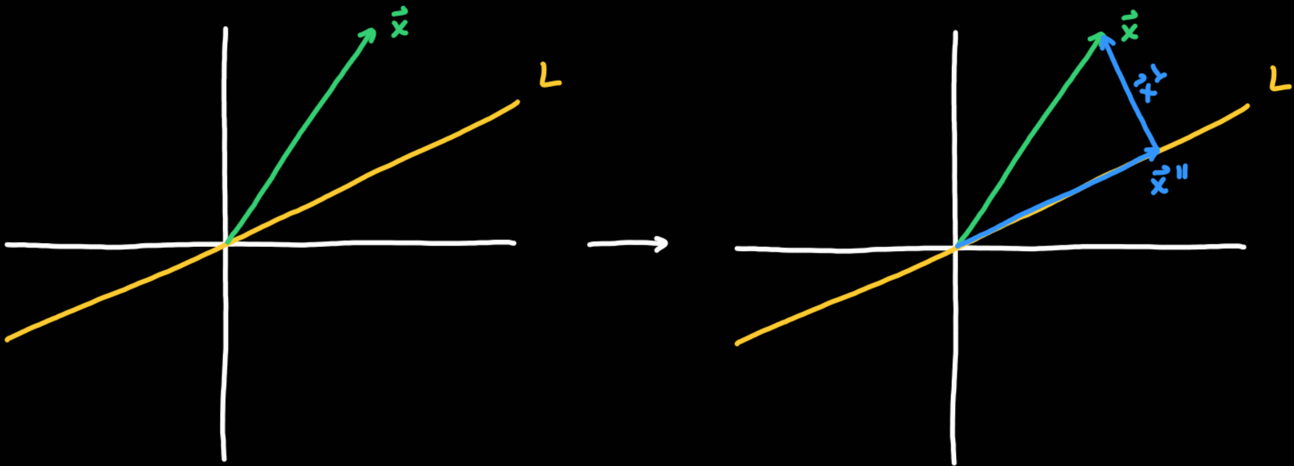
Orthogonal projections.

Given $\vec{x}$ in $\mathbb{R}^2$ and L a line through the origin, we can decompose:

$\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$ with $\vec{x}^{\parallel}$ parallel to L and $\vec{x}^{\perp}$ perpendicular to L.

We call $\vec{x}^{\parallel}$ the <u>orthogonal projection</u> of $\vec{x}$ onto L, denoted $\text{proj}_L(\vec{x})$.

We have $\vec{x}^{\parallel} = \text{proj}_L(\vec{x}) = (\vec{x} \cdot \vec{u})\vec{u}$ where $\vec{u}$ is a unit vector parallel to L.

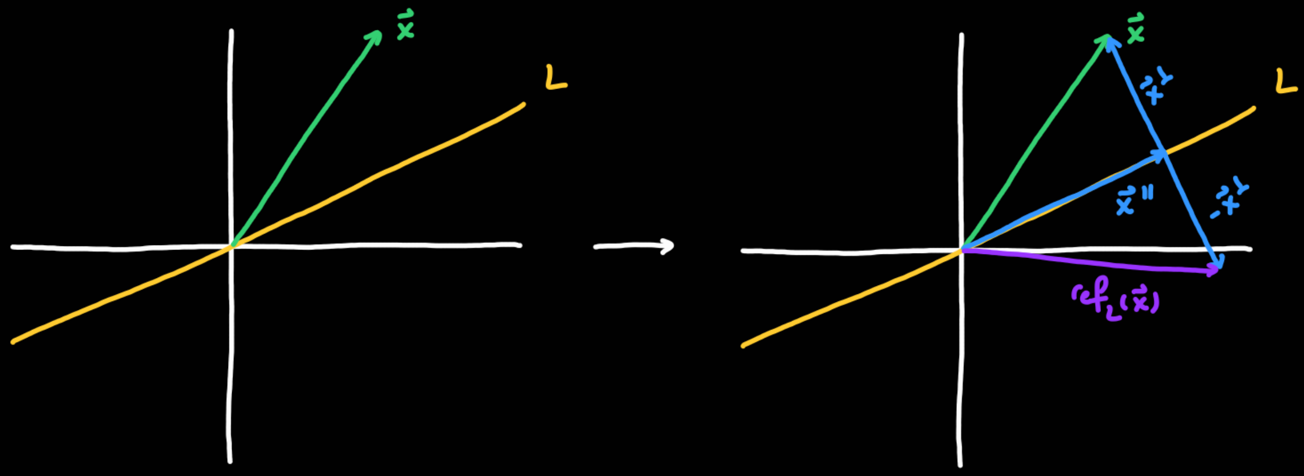


If $\vec{u}$ is non-zero and parallel to L, the associated matrix is $\frac{1}{u_1^2+u_2^2}\begin{bmatrix} u_1^2 & u_1u_2 \\ u_1u_2 & u_2^2 \end{bmatrix}$.

Reflections.

Given $\vec{x}$ in $\mathbb{R}^2$ and L a line through the origin, the <u>reflection</u> of $\vec{x}$ onto L is

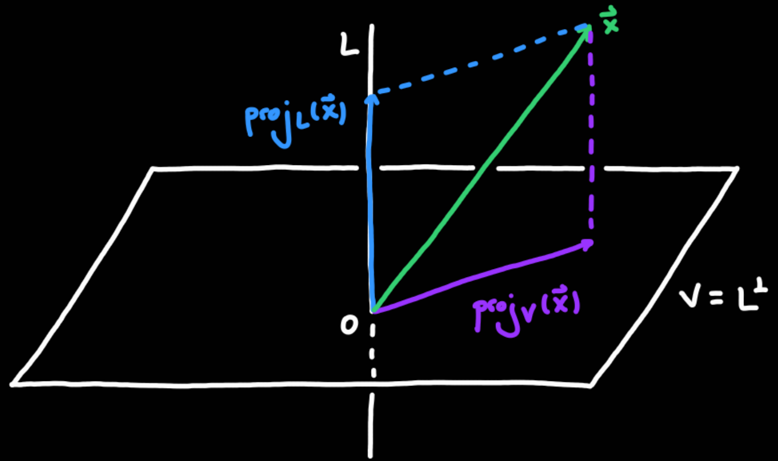
$\text{ref}_L(\vec{x}) = \vec{x}^{\parallel} - \vec{x}^{\perp}$.



If $\vec{u}$ is unitary and parallel to L, the associated matrix is $\begin{bmatrix} 2u_1^2-1 & 2u_1u_2 \\ 2u_1u_2 & 2u_2^2-1 \end{bmatrix}$.

A linear transformation is a reflection if and only if its associated matrix has the form $\begin{bmatrix} a & b \\ b & -a \end{bmatrix}$ with $a^2+b^2=1$.

We can do this same type of decompositions in higher dimensions! For $\mathbb{R}^3$, we have:



So given $\vec{x}$ in $\mathbb{R}^3$ and L a line through the origin, we can decompose $\vec{x} = \text{proj}_L(\vec{x}) + \text{proj}_V(\vec{x})$ where $\text{proj}_L(\vec{x})$ is the orthogonal projection of $\vec{x}$ onto L and $\text{proj}_V(\vec{x})$ is the projection of $\vec{x}$ onto V, the plane through the origin perpendicular to L. We have: (why? Read them on the picture!)

(i) $\text{proj}_L(\vec{x}) = (\vec{x}\cdot\vec{u})\vec{u}$

(ii) $\text{proj}_V(\vec{x}) = \vec{x} - \text{proj}_L(\vec{x})$

(iii) $\text{ref}_L(\vec{x}) = \text{proj}_L(\vec{x}) - \text{proj}_V(\vec{x})$

(iv) $\text{ref}_V(\vec{x}) = \text{proj}_V(\vec{x}) - \text{proj}_L(\vec{x})$

Example:

Let V be the plane defined by $2x_1 + x_2 - 2x_3 = 0$ and $\vec{x} = \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix}$. A vector perpendicular to V is $\vec{v} = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$, giving the unit vector $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{\sqrt{2^2+1^2+2^2}} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$, which is still perpendicular to V. Now:

(i) $\text{proj}_L(\vec{x}) = (\vec{x}\cdot\vec{u})\vec{u} = \left(5\cdot\frac{2}{3} + 4\cdot\frac{1}{3} + (-2)\cdot\frac{(-2)}{3}\right)\cdot\begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ -4 \end{bmatrix}$.

(ii) $\text{proj}_V(\vec{x}) = x - \text{proj}_L(\vec{x}) = \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix} - \begin{bmatrix} 4 \\ 2 \\ -4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$.

(iii) $\text{ref}_L(\vec{x}) = \text{proj}_L(\vec{x}) - \text{proj}_V(\vec{x}) = \begin{bmatrix} 4 \\ 2 \\ -4 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}$.

(iv) $\text{ref}_V(\vec{x}) = \text{proj}_V(\vec{x}) - \text{proj}_L(\vec{x}) = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 4 \\ 2 \\ -4 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 6 \end{bmatrix}$.

Rotations.

A linear transformation is a rotation if and only if its associated matrix has the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ with $a^2+b^2=1$.

Example: To do a rotation combined with a scaling, first do a rotation, then do a scaling. This is the same as first doing a scaling, and then a rotation.

How do we deal with consecutive linear transformations? If T is given by $\begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix}$ and S is given by $\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$, and we would like to find $\vec{z} = T(S(\vec{x}))$, we do this in two steps. Call $\vec{y} = S(\vec{x})$, then $\vec{z} = T(\vec{x})$, and these two equations are:

$$\begin{aligned} y_1 &= x_1 + 2x_2 \\ y_2 &= 3x_1 + 5x_2 \end{aligned} \quad \text{and} \quad \begin{aligned} z_1 &= 6y_1 + 7y_2 \\ z_2 &= 8y_1 + 9y_2 \end{aligned} \quad \text{so} \quad \begin{aligned} z_1 &= 27x_1 + 47x_2 . \\ z_2 &= 35x_1 + 61x_2 \end{aligned}$$

This should mean that $\vec{z} = TS(\vec{x})$ is given by $\begin{bmatrix} 27 & 47 \\ 35 & 61 \end{bmatrix}$. This should be the product of the matrices T and S, namely $\begin{bmatrix} 6 & 7 \\ 8 & 9 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$.

<u>Matrix multiplication</u>:

Let B be an $n \times p$ matrix and A a $q \times m$ matrix. If (and only if) $p = q$ the <u>product</u> BA is the matrix of the linear transformation $T(\vec{x}) = B(A\vec{x})$, and it is an $n \times m$ matrix.

<u>Theorem</u>: Let B be an $n \times p$ matrix and A a $p \times m$ matrix. Then:

(i) $$BA = B\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ B\vec{v}_1 & \cdots & B\vec{v}_m \\ | & & | \end{bmatrix}.$$

(ii) $$C = BA = \begin{bmatrix} - & \vec{w}_1 & - \\ & \vdots & \\ - & \vec{w}_n & - \end{bmatrix}\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}$$ has entries $c_{ij} = \vec{w}_i \cdot \vec{v}_j = \sum_{k=1}^{p} b_{ik} a_{kj}$.

<u>Example</u>: Matrix multiplication is not commutative:

$$\begin{bmatrix} 1 & 2 \\ \end{bmatrix}\begin{bmatrix} 1 & 1 \\ \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ \end{bmatrix} \quad \text{but} \quad \begin{bmatrix} 1 & 1 \\ \end{bmatrix}\begin{bmatrix} 1 & 2 \\ \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ \end{bmatrix}.$$

$\begin{bmatrix} 3 & 4 \end{bmatrix}\begin{bmatrix} 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 3 & 7 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} 3 & 4 \end{bmatrix} \quad \begin{bmatrix} 3 & 4 \end{bmatrix}$

<u>Algebraic rules:</u>

(i) If A is an $n \times m$ matrix then: $AI_m = I_n A = A$.

(ii) Matrix multiplication is associative: $(AB)C = A(BC)$.

(iii) Matrix multiplication distributes over matrix addition:

$$(A+B)C = AC + BC \quad \text{and} \quad A(B+C) = AB + AC.$$

(iv) Multiplication by scalars can be factored out: $(kA)B = A(kB) = k(AB)$.

A function T from X to Y is called <u>invertible</u> if for each y in Y there is a unique x in X with $T(x) = y$. The <u>inverse</u> of T, denoted T^{-1}, is a function from Y to X given by $T^{-1}(y) = x$ (it assigns to each y the x such that $T(x) = y$).

A function T has inverse L if and only if:

$$T(L(y)) = y \text{ for all } y \text{ in } Y, \text{ and } L(T(x)) = x \text{ for all } x \text{ in } X.$$

A square matrix A is said to be <u>invertible</u> if the linear transformation $\vec{y} = T(\vec{x}) = A\vec{x}$ is invertible. In this case, T^{-1} will also be a linear transformation, and we denote by A^{-1} its associated matrix: $\vec{x} = T^{-1}(\vec{y}) = A^{-1}\vec{y}$.

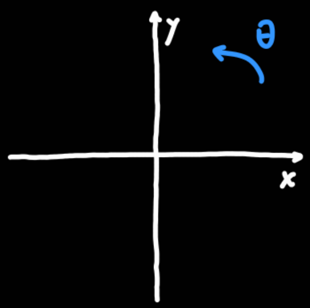
<u>Theorem:</u> Let A be an $n \times n$ matrix.

(i) A is invertible if and only if $\text{rank}(A) = n$, if and only if $\text{rref}(A) = I_n$.

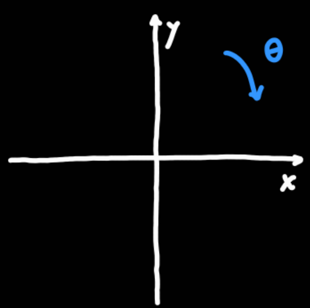
(ii) Let $\vec{b}$ be a vector in $\mathbb{R}^n$. If A is invertible then $A\vec{x} = \vec{b}$ has the unique solution $\vec{x} = A^{-1}\vec{b}$. If A is <u>not</u> invertible then $A\vec{x} = \vec{b}$ has zero or infinitely many solutions.

<u>Example:</u> Let A be an $n \times n$ matrix. The system $A\vec{x} = \vec{0}$ has $\vec{x} = \vec{0}$ as a solution. If A is invertible, this is the only solution. If A is not invertible, there are infinitely many solutions.

<u>Example:</u> Rotation by an angle θ counterclockwise is an invertible transformation with inverse rotation by an angle θ clockwise.



is given by $\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = A.$



is given by $\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} = A^{-1}!$

<u>Theorem:</u> To find the inverse of an $n \times n$ matrix A, compute $\text{rref}([A \mid I_n])$.

(i) If $\text{rref}([A \mid I_n]) = [I_n \mid B]$ then A is invertible with $A^{-1} = B$.

(ii) If $\mathrm{rref}([A \mid I_n]) \neq [I_n \mid B]$ then A is not invertible.

Example: The matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{bmatrix}$ is invertible because:

$$\mathrm{rref}\left(\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ 3 & 8 & 2 & 0 & 0 & 1 \end{array}\right]\right) = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 10 & -6 & 1 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -7 & 5 & 1 \end{array}\right], \text{ so } A^{-1} = \begin{bmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & 1 \end{bmatrix}.$$

Theorem: If A, B are two invertible $n \times n$ matrices then:

(i) $AA^{-1} = I_n$ and $A^{-1}A = I_n$,

(ii) BA is invertible with inverse $(BA)^{-1} = A^{-1}B^{-1}$.

Example:

(i) $$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{bmatrix}\begin{bmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{bmatrix}.$$

(ii) $$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Theorem: If A, B are two $n \times n$ matrices such that $BA = I_n$ then:

(i) A and B are both invertible,

(ii) $A^{-1} = B$ and $B^{-1} = A$,

(iii) $AB = I_n$.

The 2×2 matrix $A = \begin{bmatrix} a & b \\ \cdot & \cdot \end{bmatrix}$ has <u>determinant</u> $\det(A) = ad - bc$. The matrix A is invertible if and only if $\det(A) \neq 0$. If A is invertible then $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Examples:

(i) For which values of k is the matrix $A = \begin{bmatrix} 1-k & 2 \\ 4 & 3-k \end{bmatrix}$ invertible?

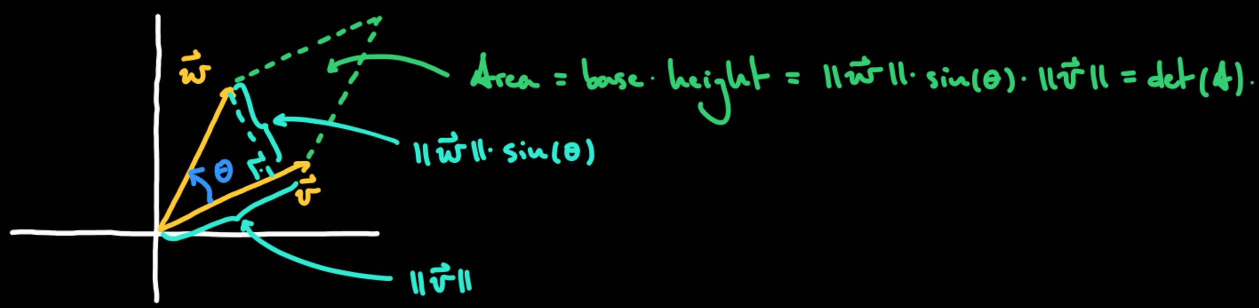
The matrix has determinant $\det(A) = (k-5)(k+1)$, so if $k \neq 5, -1$ then $\det(A) \neq 0$ and A is invertible.

(ii) Let A be a matrix representing the reflection about a line L passing through the origin. Then $A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$ with $a^2+b^2=1$, so $\det(A) = -a^2-b^2 = -1$ and the inverse is $A^{-1} = \frac{1}{-1}\begin{bmatrix} -a & -b \\ -b & a \end{bmatrix} = \begin{bmatrix} a & b \\ b & -a \end{bmatrix} = A$. That is, a reflection is its own inverse.

For $A = [\vec{v}\ \vec{w}]$ with $\vec{v}, \vec{w}$ in $\mathbb{R}^2$ we have:

$$\det(A) = \|\vec{v}\| \cdot \|\vec{w}\| \cdot \sin(\theta) \quad \text{with } \theta \text{ the angle between } \vec{v} \text{ and } \vec{w}.$$

This determinant is the area of the parallelogram spanned by $\vec{v}$ and $\vec{w}$.



2. Subspaces of $\mathbb{R}^n$ (Chapter 3)

The <u>image</u> of a function T is the set of values $\text{im}(T)$ that the function takes.

$$\text{im}(T) = \{ T(x) \mid x \in X \} = \{ y \in Y \mid y = T(x) \text{ for some } x \in X \}.$$

<u>Example:</u> Find the image of the following linear transformations:

1. $T: \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(\vec{x}) = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \vec{x}$.

$$T(\vec{x}) = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 6 \end{bmatrix} = (x_1 + 3x_2) \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

So the image of T is the scalar multiples of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$:

$$\text{im}(T) = \{ k \begin{bmatrix} 1 \\ 2 \end{bmatrix} \mid k \in \mathbb{R} \}.$$

2. $T: \mathbb{R}^3 \to \mathbb{R}^2$ given by $T(\vec{x}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \vec{x}$.

This is an orthogonal projection onto the x_1-x_2-plane: $T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$, so the image consists of the x_1-x_2-plane:

$$\text{im}(T) = \{ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R} \}.$$

Let $\vec{v}_1, \dots, \vec{v}_m$ be vectors in $\mathbb{R}^n$, the set of all linear combinations of $\vec{v}_1, \dots, \vec{v}_m$ is called their <u>span</u>:

$$\text{span}(\vec{v}_1, \dots, \vec{v}_m) = \{ c_1 \vec{v}_1 + \dots + c_m \vec{v}_m \mid c_1, \dots, c_m \in \mathbb{R} \}.$$

Note that the image of the linear transformation $T(\vec{x}) = A\vec{x}$ is the span of the column vectors of A.

<u>Theorem:</u> Let $T: \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation, then: (why? What does T

linear imply ?)

(i) The zero vector $\vec{0} \in \mathbb{R}^n$ is in the image of T.

(ii) The image of T is closed under addition.

(iii) The image of T is closed under scalar multiplication.

<u>Example:</u> Consider $T: \mathbb{R}^2 \to \mathbb{R}^2$ given by $T(\vec{x}) = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \vec{x}$. Then:

(i) $T(\vec{0}) = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

(ii) $a \begin{bmatrix} 1 \\ 2 \end{bmatrix} + b \begin{bmatrix} 1 \\ 2 \end{bmatrix} = (a+b) \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

(iii) $k \left(a \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right) = (ka) \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

The <u>kernel</u> of a function T is the set of values $\ker(T)$ that the function takes to $\vec{0}$.

$\ker(T) = \{ \vec{x} \in \mathbb{R}^m \mid T(\vec{x}) = \vec{0} \}$, the solutions of the equation $T(\vec{x}) = \vec{0}$.

<u>Example:</u> Find the kernel of the linear transformation $T: \mathbb{R}^3 \to \mathbb{R}^2$ given by $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$.

We solve the equation $T(\vec{x}) = \vec{0}$, which is the linear system:

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 + 2x_3 + 3x_3 &= 0 \end{aligned}$$, with augmented matrix $\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & 2 & 3 & 0 \end{array}\right]$.

This matrix has reduced form $\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right]$, so by setting $x_1 = t$ the free variable we have solutions $\vec{x} = \begin{bmatrix} t \\ -2t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$. Thus the kernel of T are scalar multiples of $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$.

$\ker(T) = \left\{ k \begin{bmatrix} \\ -2 \\ 1 \end{bmatrix} \mid k \in \mathbb{R} \right\}$.

<u>Example:</u> Let T be an invertible linear transformation. If $\vec{x} \in \ker(T)$, then $T(\vec{x}) = \vec{0}$

so $A\vec{x} = \vec{0}$ so $\vec{x} = A^{-1}A\vec{x} = A^{-1}\vec{0} = \vec{0}$. Hence $\ker(T) = \{\vec{0}\}$, often denoted $\ker(T) = 0$.

<u>Example:</u> Let A be an $n \times m$ matrix with $\ker(A) = 0$. Then, the system $A\vec{x} = \vec{0}$ has exactly one solution, and thus there are no free variables. In particular, all variables are leading, so $\text{rank}(A) = m$. (is the converse true? Why?)

<u>Theorem:</u> Let $T: \mathbb{R}^m \longrightarrow \mathbb{R}^n$ be a linear transformation, then: (why? What does T linear imply?)

(i) The zero vector $\vec{0} \in \mathbb{R}^m$ is in the kernel of T.

(ii) The kernel of T is closed under addition.

(iii) The kernel of T is closed under scalar multiplication.

<u>Theorem:</u> Let A be an $n \times m$ matrix.

(i) We have $\ker(A) = 0$ if and only if $\text{rank}(A) = m$.

(ii) If $\ker(A) = 0$ then $m \leq n$.

(iii) If $m > n$ then there are non-zero vectors in the kernel of A.

(iv) Let $n = m$, so A is a square matrix. We have $\ker(A) = 0$ if and only if A is invertible.

<u>Recall:</u> Let A be an $n \times n$ matrix. The following statements are equivalent:

(i) A is invertible.

(ii) The equation $A\vec{x} = \vec{b}$ has a unique solution for each $\vec{b} \in \mathbb{R}^n$.

(iii) $\operatorname{rref}(A) = I_n$.

(iv) $\operatorname{rank}(A) = n$.

(v) $\operatorname{im}(A) = \mathbb{R}^n$.

(vi) $\ker(A) = \{\vec{0}\}$.

A subset W of the vector space $\mathbb{R}^n$ is called a <u>linear subspace</u> of $\mathbb{R}^n$ if:

(i) The zero vector $\vec{0}$ is in W, and

(ii) W is closed under addition, and

(iii) W is closed under scalar multiplication.

In particular, W is closed under linear combinations.

<u>Example:</u> Let $T: \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. Then $\operatorname{im}(T)$ is a subspace of $\mathbb{R}^n$ and $\ker(T)$ is a subspace of $\mathbb{R}^m$.

<u>Example:</u>

1. The set $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \,\middle|\, x \geq 0, y \geq 0 \right\}$ is not a subspace of $\mathbb{R}^2$.

2. The only subspaces of $\mathbb{R}^2$ are:

(a) $\mathbb{R}^2$,

(b) $\{\vec{0}\}$,

(c) Any line L passing through the origin.

3. Let V be a plane in $\mathbb{R}^3$ given by the equation $x+2y+3z=0$.

(a) Find a matrix A such that $\ker(A) = V$.

We want a matrix that inputs a vector in $\mathbb{R}^3$ and whose kernel is a plane, that is, the equation $A\vec{x}=\vec{0}$ is the defining equation of V. Since V has only one defining equation, $A\vec{x}=\vec{0}$ has to already be one equation, so $\vec{0}=[0]$ is in $\mathbb{R}$. Thus A inputs a vector in $\mathbb{R}^3$ and outputs a vector in $\mathbb{R}$, so A is a 1×3 matrix. Now:

$$0 = A\vec{x} = \begin{bmatrix} a & b & c \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = ax+by+cz$$

has to be the equation $x+2y+3z=0$ defining V. Thus $a=1$, $b=2$, $c=3$ and hence $A = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$.

(b) Find a matrix B such that $\operatorname{im}(B) = V$.

We know that the image of a matrix is the span of its columns, so if we describe V as the span of two vectors, the matrix B will have two columns, each column will be one of those vectors. To find two non parallel vectors, we set $z=0$ and obtain that $\begin{bmatrix}-2\\1\\0\end{bmatrix}$ is in V, and we set $y=0$ and obtain that $\begin{bmatrix}-3\\0\\1\end{bmatrix}$ is in V. Since their dot product is 6, which is not zero, these vectors are not parallel, they are both in V, so they span V. Thus $B=\begin{bmatrix}-2 & -3\\1 & 1\\0 & 0\end{bmatrix}$.

<u>Example:</u> Find vectors in $\mathbb{R}^3$ that span the image of $A=\begin{bmatrix}1 & 2 & 1 & 2\\1 & 2 & 2 & 3\\1 & 2 & 3 & 4\end{bmatrix}$. Find the smallest number of vectors needed to span the image of A.

We know that the image of A is spanned by the column vectors of A:

$$\vec{v}_1=\begin{bmatrix}1\\1\\1\end{bmatrix},\ \vec{v}_2=\begin{bmatrix}2\\2\\2\end{bmatrix},\ \vec{v}_3=\begin{bmatrix}1\\2\\3\end{bmatrix},\ \vec{v}_4=\begin{bmatrix}2\\3\\4\end{bmatrix},\quad \operatorname{im}(A)=\operatorname{span}(\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4).$$

However, since $\vec{v}_2=2\cdot\vec{v}_1$ and $\vec{v}_4=\vec{v}_1+\vec{v}_3$, we have that $\vec{v}_2$ and $\vec{v}_4$ do not contribute to the span of $\vec{v}_1$ and $\vec{v}_3$. Namely, let $\vec{v}\in\operatorname{span}(\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4)$, then:

$$\vec{v}=c_1\vec{v}_1+c_2\vec{v}_2+c_3\vec{v}_3+c_4\vec{v}_4=(c_1+2c_2+c_4)\vec{v}_1+(c_3+c_4)\vec{v}_3$$

for c_1,c_2,c_3,c_4 real numbers, so $\vec{v}\in\operatorname{span}(\vec{v}_1,\vec{v}_3)$. Thus:

$$\operatorname{im}(A)=\operatorname{span}(\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4)=\operatorname{span}(\vec{v}_1,\vec{v}_3)$$

and since $\vec{v}_1\cdot\vec{v}_3=6$ they are not parallel, and two is the minimum number of vectors needed to span the image of A.

Let $\vec{v}_1, \dots, \vec{v}_m$ be vectors in $\mathbb{R}^n$. We say that a vector $\vec{v}_i$ is *redundant* if $\vec{v}_i$ is a linear combination of $\vec{v}_1, \dots, \vec{v}_{i-1}$. We say that $\vec{v}_1$ is *redundant* if $\vec{v}_1 = \vec{0}$. We say that the vectors $\vec{v}_1, \dots, \vec{v}_m$ are *linearly independent* if none of them is redundant. If at least one of them is redundant, we call them *linearly dependent*. We say that the vectors $\vec{v}_1, \dots, \vec{v}_m$ form a *basis* of a subspace V of $\mathbb{R}^n$ if they span V and are linearly independent.

Example: To construct a basis of the image of a matrix A we only need to list the column vectors of A and remove the redundant vectors from the list.

Example: We can identify linearly independence if enough entries are zero:

1. $\vec{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 9 \\ 7 \\ 3 \end{bmatrix}$ are linearly independent.
2. $\vec{v}_1 = \begin{bmatrix} 7 \\ 0 \\ 4 \\ 0 \\ 1 \\ 9 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 6 \\ 0 \\ 7 \\ 1 \\ 4 \\ 8 \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 5 \\ 0 \\ 6 \\ 2 \\ 3 \\ 1 \\ 7 \end{bmatrix}, \vec{v}_4 = \begin{bmatrix} 4 \\ 5 \\ 3 \\ 3 \\ 2 \\ 2 \\ 4 \end{bmatrix}$ are linearly independent.

Let $\vec{v}_1, \dots, \vec{v}_m$ be vectors in $\mathbb{R}^n$. An equation of the form:

$$c_1\vec{v}_1 + \cdots + c_m\vec{v}_m = \vec{0}$$

is called a *linear relation* among $\vec{v}_1, \dots, \vec{v}_m$. If $c_1 = \cdots = c_m$, the relation is called *trivial*.

If at least one c_i is non zero, the relation is <u>nontrivial</u>.

Theorem: The vectors $\vec{v}_1, \ldots, \vec{v}_m$ in $\mathbb{R}^n$ are linearly dependent if and only if there is at least one nontrivial relation among them.

Example: Let A be an $n \times m$ matrix with linearly independent columns. Then the equation

$A\vec{x} = \vec{0}$ with $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}$ gives a relation $x_1\vec{v}_1 + \cdots + x_m\vec{v}_m = \vec{0}$, so by

linear independence this relation must be trivial, so $x_1 = \cdots = x_m = 0$, so $\ker(A) = 0$.

More generally, the vectors in the kernel of an $n \times m$ matrix A correspond to the linear relations among the column vectors of A, since writing $A\vec{x} = \vec{0}$ with $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}$ means that $x_1\vec{v}_1 + \cdots + x_m\vec{v}_m = \vec{0}$. In other words, the column vectors of A are linearly independent if and only if $\ker(A) = 0$, which happens if and only if $\operatorname{rank}(A) = m$.

Example: The columns of the matrix $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ are linearly dependent since:

$$1 \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - 2 \cdot \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + 1 \cdot \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ and thus } \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \text{ is in } \ker(A).$$

Recall: Let $\vec{v}_1, \ldots, \vec{v}_m$ be vectors in $\mathbb{R}^n$. The following are equivalent:

(i) The vectors $\vec{v}_1, \ldots, \vec{v}_m$ are linearly independent.

(ii) None of the vectors $\vec{v}_1, \ldots, \vec{v}_m$ is redundant.

(iii) None of the vectors $\vec{v}_1, \ldots, \vec{v}_m$ is a linear combination of the others.

(iv) There is only one trivial relation among the vectors $\vec{v}_1, \dots, \vec{v}_m$.

(v) $\ker\left(\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}\right) = 0$.

(vi) $\operatorname{rank}\left(\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}\right) = m$.

<u>Theorem:</u> Let V be a subspace of $\mathbb{R}^n$, let $\vec{v}_1, \dots, \vec{v}_m$ be vectors in V. They form a basis of V if and only if every vector $\vec{v}$ in V can be expressed uniquely as a linear combination:

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m .$$

We call $c_1, \dots, c_m$ the <u>coordinates</u> of $\vec{v}$ with respect to the basis $\vec{v}_1, \dots, \vec{v}_m$.

<u>Theorem:</u>

(i) A spanning set is larger than or equal in size to a linearly independent set.

(ii) Any two basis of the same subspace have the same number of elements.

The number $\dim(V)$ of vectors in a basis of a subspace V is called the <u>dimension</u> of V.

<u>Theorem:</u> Let V be a subspace of $\mathbb{R}^n$ with $\dim(V) = m$.

(i) There are at most m linearly independent vectors in V.

(ii) We need at least m vectors to span V.

(iii) If m vectors in V are linearly independent, they form a basis of V.

(iv) If m vectors in V span V, they form a basis of V.

There are strong relations between the size of a matrix, its rank, and the dimensions of its kernel and image.

<u>Theorem</u>: Let A be a matrix, then $\dim(\text{im}(A)) = \text{rank}(A)$.

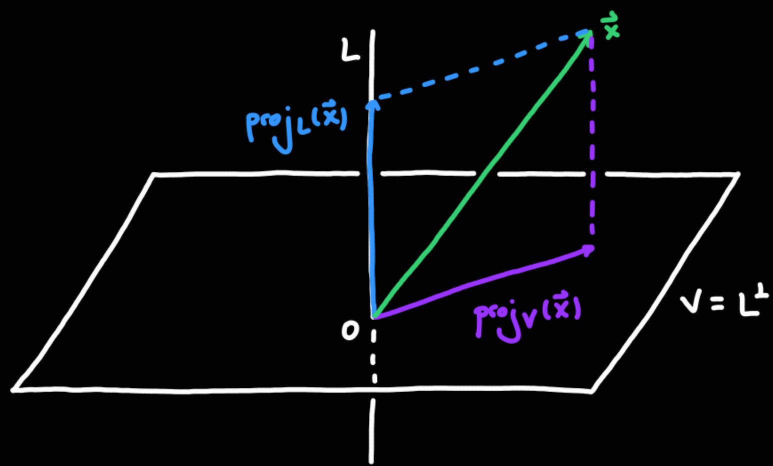
<u>Theorem</u>: (Rank-Nullity) Let A be an $n \times m$ matrix, then:

$$\dim(\ker(A)) + \dim(\text{im}(A)) = m.$$

We call the dimension of $\ker(A)$ the <u>nullity</u> of A. Then:

$$(\text{nullity of } A) + (\text{rank of } A) = m.$$

<u>Example</u>: Let T be the orthogonal projection onto a plane V in $\mathbb{R}^3$.



Here we have $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ so it is given by a 2×3 matrix so $m = 3$, also $\text{im}(T) = V$ so $\dim(\text{im}(T)) = 2$, and $\ker(T) = L$ so $\dim(\ker(T)) = 1$. Clearly $1 + 2 = 3$.

<u>Theorem</u>: Let $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}$ be an $n \times m$ matrix with $\text{rref}(A) = \begin{bmatrix} | & & | \\ \vec{w}_1 & \cdots & \vec{w}_m \\ | & & | \end{bmatrix}$.

(i) To construct a basis of the image of A, pick the column vectors of A that correspond to the columns of $\operatorname{rref}(A)$ containing the leading 1's.

(ii) The column vectors of A that correspond to the columns of $\operatorname{rref}(A)$ that do not contain leading 1's are redundant, and they form a basis of $\ker(A)$.

(iii) Suppose that column i of $\operatorname{rref}(A)$ does not contain a leading 1, let $w_{j_1}, \dots, w_{j_r}$ be the entries of $\vec{w}_i$ that have a leading 1 to its left, say in columns $c_1, \dots, c_r$ respectively. Then we have the relation: $\vec{v}_i = w_{j_1}\vec{v}_{c_1} + \cdots + w_{j_r}\vec{v}_{c_r}$.

In particular, this last relation means that the vector with entry 1 in position i and entries $-w_{j_1}, \dots, -w_{j_r}$ in positions $c_1, \dots, c_r$ is in the kernel of A, and all these vectors form a basis of $\ker(A)$.

<u>Example:</u> Find bases of the image and kernel of $A = \begin{bmatrix} 1 & 2 & 0 & 1 & 2 \\ 1 & 2 & 0 & 2 & 3 \\ 1 & 2 & 0 & 3 & 4 \\ 1 & 2 & 0 & 4 & 5 \end{bmatrix}$.

We have $\operatorname{rref}(A) = \begin{bmatrix} 1 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, so seeing $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}$ we have $\vec{v}_1$ and $\vec{v}_4$ as basis of the image:

$$\operatorname{im}(A) = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \right\}.$$

Moreover, since columns 2, 3, 5 do not have leading 1's, vectors $\vec{v}_2, \vec{v}_3, \vec{v}_5$ are redundant.

We can read from $\operatorname{rref}(A)$ that:

$$\vec{v}_2 = 2\vec{v}_1 \quad \text{so} \quad -2\vec{v}_1 + \vec{v}_2 = \vec{0} \quad \text{so} \quad \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ is in } \ker(A).$$

$$\vec{v}_3 = 0 \cdot \vec{v}_1 = \vec{0} \quad \text{so} \quad \vec{v}_3 = \vec{0} \quad \text{so} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ is in } \ker(A).$$

$$\vec{v}_5 = \vec{v}_1 + \vec{v}_4 \quad \text{so} \quad -\vec{v}_1 - \vec{v}_4 + \vec{v}_5 = \vec{0} \quad \text{so} \quad \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} \text{ is in } \ker(A).$$

In fact, these three vectors form a basis of $\ker(A)$:

$$\ker(A) = \operatorname{span}\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

So far, the only basis of $\mathbb{R}^n$ that we have seen is the canonical one, namely:

$$\mathbb{R}^n = \operatorname{span}\left\{ \vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \vec{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right\}.$$

However, there are many more.

<u>Theorem:</u> The vectors $\vec{v}_1, \dots, \vec{v}_n$ of $\mathbb{R}^n$ form a basis of $\mathbb{R}^n$ if and only if the matrix

$$\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix} \text{ is invertible.}$$

<u>Example:</u> For which values of the constant k do the vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ k \\ k^2 \end{bmatrix}$ form a basis of $\mathbb{R}^2$? It suffices to examine the matrix $\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & k \\ 1 & 1 & k^2 \end{bmatrix}$, and to determine when it is invertible. Since $\operatorname{rref}\left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & k \\ 1 & 1 & k^2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & (1-k)/2 \\ 0 & 0 & k^2-1 \end{bmatrix}$, the only thing that we need for this reduced matrix to be further reducible to I_3 is that $k^2 - 1 \neq 0$, that is $k \neq 1, -1$.

Coordinates in a subspace of $\mathbb{R}^n$.

Let $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_m\}$ be a basis of a subspace V of $\mathbb{R}^n$. Any $\vec{x}$ in V can be written as $\vec{x} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$ for some real scalars $c_1, \dots, c_m$, called the <u>$\mathcal{B}$-coordinates</u> of $\vec{x}$. The vector $\begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}$ is called the <u>$\mathcal{B}$-coordinate vector</u> of $\vec{x}$, denoted $[\vec{x}]_{\mathcal{B}}$.

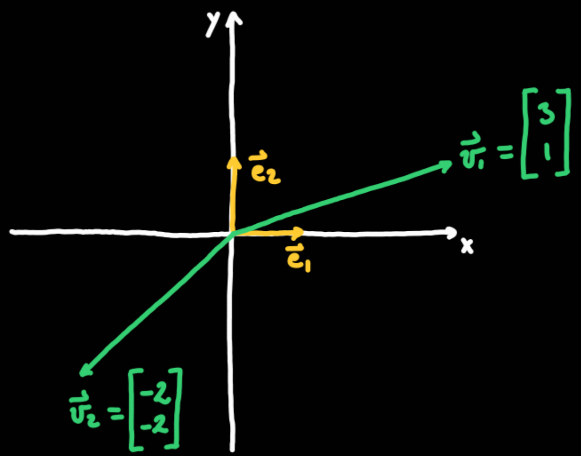
<u>Example:</u> Note that:

$$\vec{x} = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}\begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = S\,[\vec{x}]_{\mathcal{B}}.$$

The matrix S inputs vectors with $\mathcal{B}$-coordinates, and outputs vectors in the standard basis $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \vec{e}_m = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$.

(later, we will call this a change of basis matrix)

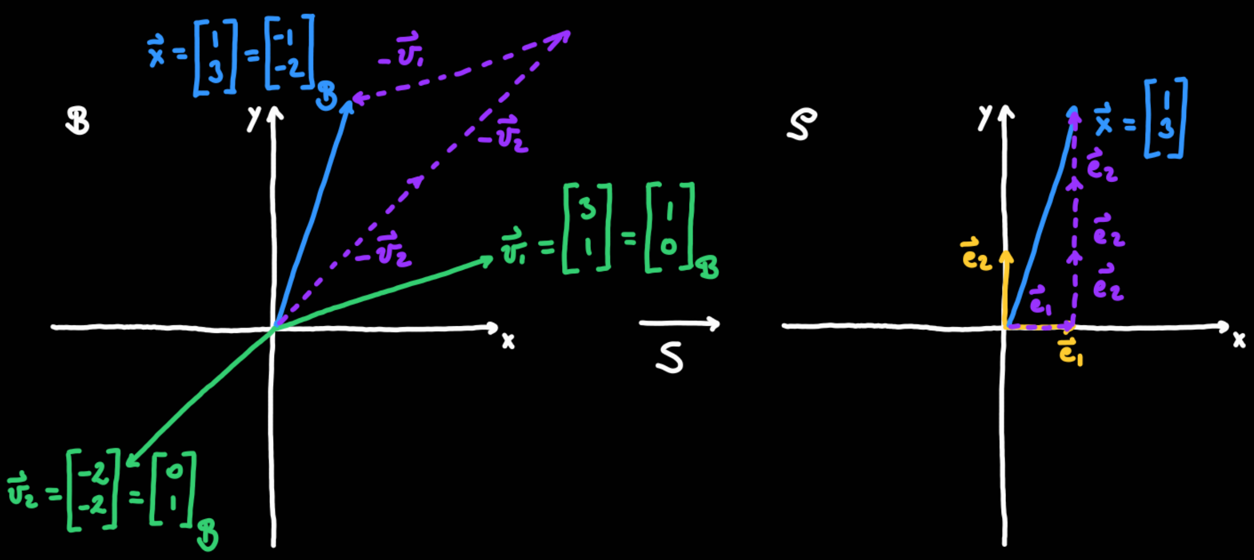
<u>Example:</u> The plane $\mathbb{R}^2$ has a basis $\mathcal{B} = \left\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \end{bmatrix}\right\}$.



The matrix $S = \begin{bmatrix} 3 & -2 \\ 1 & -2 \end{bmatrix}$ takes $\begin{bmatrix} 1 \\ 0 \end{bmatrix}_{\mathcal{B}}$ to $\vec{v}_1$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}_{\mathcal{B}}$ to $\vec{v}_2$. It is a transformation

$$S : \mathbb{R}^2_{\mathcal{B}} \longrightarrow \mathbb{R}^2_{\mathcal{S}}.$$

The standard basis of $\mathbb{R}^2$ is $\mathcal{S} = \left\{\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$. When we write a vector $\vec{x}$ as $\vec{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, we are implicitly writing $[\vec{x}]_{\mathcal{S}} = 1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in the target of the linear transformation S. When we write $\vec{x}$ as $[\vec{x}]_{\mathcal{B}} = -1 \cdot \begin{bmatrix} 3 \\ 1 \end{bmatrix} - 2 \cdot \begin{bmatrix} -2 \\ -2 \end{bmatrix}$, we are now in the source of the linear transformation S.



The matrix S is sending a vector in the basis $\mathcal{B}$ to a vector in the basis $\mathcal{S}$.

When we are given a vector in the basis $\mathcal{S}$ and we are asked to find its coordinates in the basis $\mathcal{B}$, we are being asked to solve the system $\vec{x} = S[\vec{x}]_{\mathcal{B}}$, where $\vec{x}$ and S are known, and $[\vec{x}]_{\mathcal{B}}$ is unknown. Since S has for columns the vectors in the basis $\mathcal{B}$, the matrix S has all columns linearly independent. If S is a square matrix, S will then be invertible, and $[\vec{x}]_{\mathcal{B}} = S^{-1}\vec{x}$.

Let $\vec{x} = \begin{bmatrix}3\\-3\end{bmatrix}$, to find $[\vec{x}]_{\mathcal{B}}$ we solve:

$\begin{bmatrix}3\\-3\end{bmatrix} = \begin{bmatrix}3 & -2\\1 & -2\end{bmatrix}\begin{bmatrix}c_1\\c_2\end{bmatrix}$, with augmented matrix $\left[\begin{array}{cc|c}3 & -2 & 3\\1 & -2 & -3\end{array}\right]$, which reduces to $\left[\begin{array}{cc|c}1 & 0 & 3\\0 & 1 & 3\end{array}\right]$, and thus $c_1 = 3$ and $c_2 = 3$, so $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix}3\\3\end{bmatrix}$.

Alternatively, we compute:

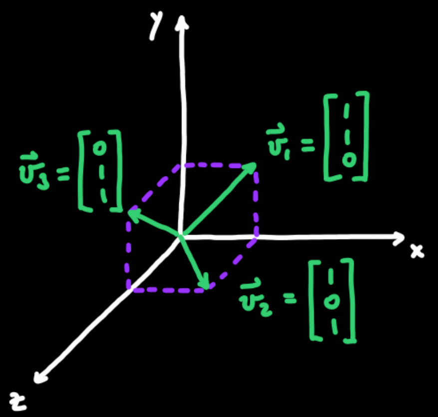
$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix}3 & -2\\1 & -2\end{bmatrix}^{-1}\begin{bmatrix}3\\-3\end{bmatrix} = \frac{1}{-4}\begin{bmatrix}-2 & 2\\-1 & 3\end{bmatrix}\begin{bmatrix}3\\-3\end{bmatrix} = \begin{bmatrix}3\\3\end{bmatrix}.$$

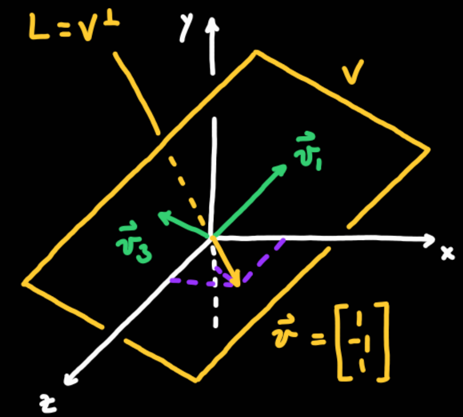
Theorem: Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation, $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ a basis of $\mathbb{R}^n$.

Then there exists a unique $n \times n$ matrix B transforming $[\vec{x}]_{\mathcal{B}}$ into $[T(\vec{x})]_{\mathcal{B}}$,

namely $[T(\vec{x})]_{\mathcal{B}} = B[\vec{x}]_{\mathcal{B}}$. Moreover: $B = \begin{bmatrix} | & & | \\ [T(\vec{v}_1)]_{\mathcal{B}} & \cdots & [T(\vec{v}_n)]_{\mathcal{B}} \\ | & & | \end{bmatrix}$.

Example: Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ be a basis of $\mathbb{R}^3$. Consider $T: \mathbb{R}^3 \to \mathbb{R}^3$ the linear transformation that projects any vector orthogonally onto the plane V spanned by $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.





To find the defining equation of V, we compute $\vec{v}$ the cross product of $\vec{v}_1$ and $\vec{v}_2$, obtaining a vector perpendicular to V, so V is given by $v_1 x + v_2 y + v_3 z = 0$. In our particular case, $\vec{v} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ so V is given by $x - y + z = 0$.

To find the matrix associated to T in the standard basis $\mathcal{S} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$, we know that $A = \begin{bmatrix} | & | & | \\ T(\vec{e}_1) & T(\vec{e}_2) & T(\vec{e}_3) \\ | & | & | \end{bmatrix}$. Recall that:

$$T(\vec{x}) = \text{proj}_V(\vec{x}) = \vec{x} - \text{proj}_L(\vec{x}) = \vec{x} - (\vec{x} \cdot \vec{u}) \cdot \vec{u}$$

with $\vec{n} = \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\-1\\1\end{bmatrix}$, so now:

$$T(\vec{e}_1) = \vec{e}_1 - \frac{1}{\sqrt{3}}\cdot\frac{1}{\sqrt{3}}\begin{bmatrix}1\\-1\\1\end{bmatrix} = \begin{bmatrix}2/3\\1/3\\-1/3\end{bmatrix},$$

$$T(\vec{e}_2) = \vec{e}_2 + \frac{1}{\sqrt{3}}\cdot\frac{1}{\sqrt{3}}\begin{bmatrix}1\\-1\\1\end{bmatrix} = \begin{bmatrix}1/3\\2/3\\1/3\end{bmatrix},$$

$$T(\vec{e}_3) = \vec{e}_3 - \frac{1}{\sqrt{3}}\cdot\frac{1}{\sqrt{3}}\begin{bmatrix}1\\-1\\1\end{bmatrix} = \begin{bmatrix}-1/3\\1/3\\2/3\end{bmatrix},$$

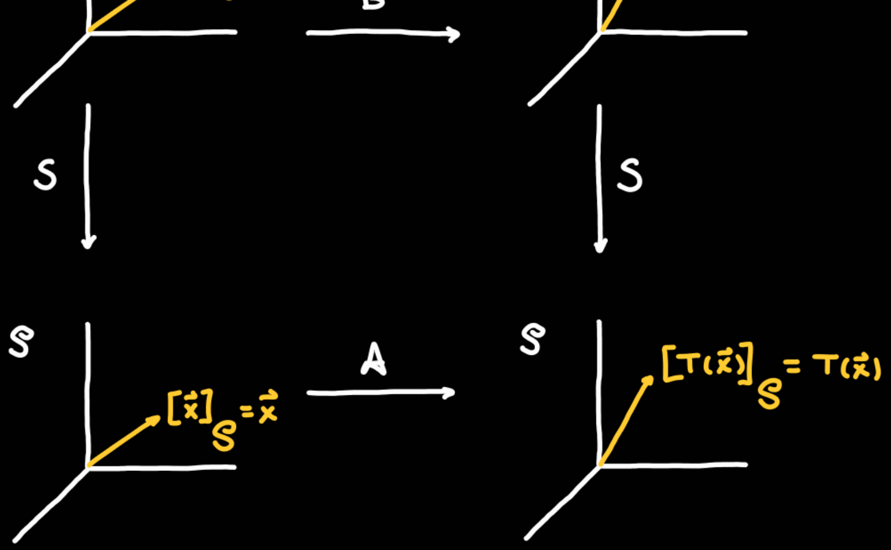
and thus $A = \begin{bmatrix}2/3 & 1/3 & -1/3\\ 1/3 & 2/3 & 1/3\\ -1/3 & 1/3 & 2/3\end{bmatrix}$.

All the above work was done over the standard basis S. If we work with the basis $\mathcal{B}$, the linear transformation $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ is projecting onto the basis vectors $\vec{v}_1$ and $\vec{v}_3$. Note that:

$\vec{v}_2 = \frac{1}{3}\vec{v}_1 + \frac{1}{3}\vec{v}_3 + \frac{2}{3}\vec{v}$, where we computed $\text{proj}_{L}(\vec{v}_2) = \frac{2}{3}\vec{v}$ as the component of $\vec{v}_2$ that gets sent to zero by T.

Thus if we write $\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \left(c_1 + \frac{c_2}{3}\right)\vec{v}_1 + \left(c_3 + \frac{c_2}{3}\right)\vec{v}_3 + \frac{2c_2}{3}\vec{v}$, we have $T(\vec{x}) = \left(c_1 + \frac{c_2}{3}\right)\vec{v}_1 + \left(c_3 + \frac{c_2}{3}\right)\vec{v}_3$. As before, we know that working over $\mathcal{B}$, the matrix associated to T should be $B = \begin{bmatrix} | & | & | \\ [T(\vec{v}_1)]_{\mathcal{B}} & [T(\vec{v}_2)]_{\mathcal{B}} & [T(\vec{v}_3)]_{\mathcal{B}} \\ | & | & | \end{bmatrix}$,

namely $B = \begin{bmatrix}1 & 1/3 & 0\\ 0 & 0 & 0\\ 0 & 1/3 & 1\end{bmatrix}$. The following is happening:



Namely:

$T(\vec{x}) = AS[\vec{x}]_B$ and $T(\vec{x}) = SB[\vec{x}]_B$ for all $\vec{x}$, so $AS = SB$

where $S = \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ is the matrix who inputs vectors in the basis $\mathcal{B}$ and outputs vectors in the basis $\mathcal{S}$. In other words:

$$\begin{tikzcd} \mathbb{R}^3, \mathcal{S} \arrow[r, "A"] \arrow[d, "S^{-1}"'] & \mathbb{R}^3, \mathcal{S} \\ \mathbb{R}^3, \mathcal{B} \arrow[r, "B"] & \mathbb{R}^3, \mathcal{B} \arrow[u, "S"'] \end{tikzcd}$$

to compute T we can work in any basis, each basis will have a matrix that is associated to T, and we can move back and forth between the different basis using the change of basis matrix.

We can check that:

$$SBS^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1/3 & 0 \\ 0 & 0 & 0 \\ 0 & 1/3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}^{-1} =$$

$$= \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1/3 & 0 \\ 0 & 0 & 0 \\ 0 & 1/3 & 1 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ -1/2 & 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 2/3 & 1/3 & -1/3 \\ 1/3 & 2/3 & 1/3 \\ -1/3 & 1/3 & 2/3 \end{bmatrix} = A.$$

__Theorem:__ Let $T: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation, $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ a basis of $\mathbb{R}^n$, A the matrix associated to T in the standard basis, B the matrix associated to T in the basis $\mathcal{B}$, and $S = \begin{bmatrix} | & & | \\ \vec{v}_1 & \dots & \vec{v}_n \\ | & & | \end{bmatrix}$. Then $AS = SB$.

Let A, B be two $n \times n$ matrices. We say that A is __similar__ to B if there exists an invertible matrix S such that $AS = SB$.

__Example:__ Let A be similar to B, then A^T is similar to B^T:

$$A^T = (SBS^{-1})^T = (S^{-1})^T B^T S^T = (S^T)^{-1} B^T S^T = R^{-1} B^T R \quad \text{with } R = S^T.$$

3. Orthogonality. (Chapter 4)