MATH 33A – WINTER 2022

Pablo S. Ocal

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

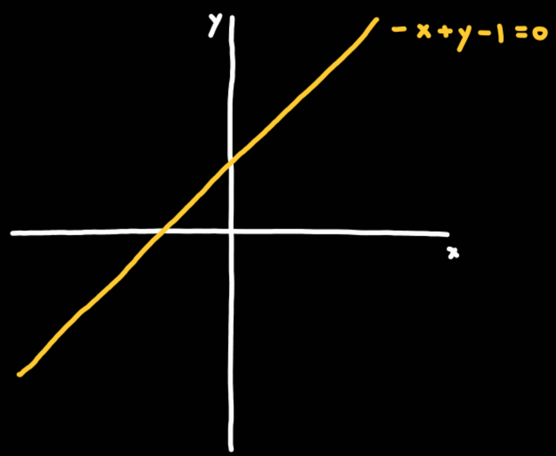
1. Introduction to Linear Algebra. (Chapter 1 and Chapter 2)

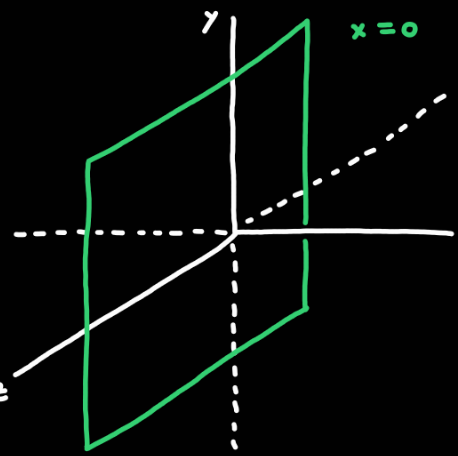
Linear algebra is the study of <u>linear equations</u> and <u>linear transformations</u>.

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



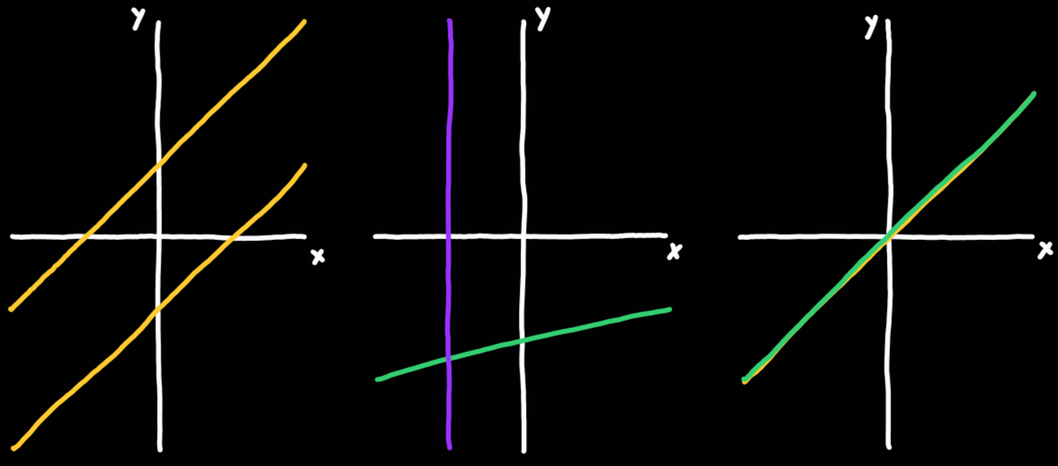


<u>Systems of linear equations</u> can have no solution, one solution, or infinitely many solutions.

$$a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n + b_1 = 0$$
$$\vdots$$

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

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.

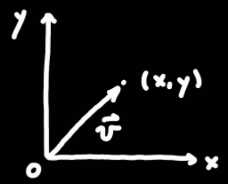
A vector is a matrix with only one column. The entries of a vector are called its components.

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

vector space.

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 augmented matrix:

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

and simplify it using three row operations: (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

$$\begin{aligned} x &= 11 \\ y &= -4 \\ z &= 3. \end{aligned}$$

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.

Example: The zero matrix is in reduced row-echelon form.

Example: 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 consistent if there is at least one solution, and inconsistent if there are no solutions.

Theorem: A linear system is inconsistent if and only if the reduced row-echelon form of its augmented matrix contains the row $[0 \cdots 0 \,1 \,|\, 1]$. 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 $\text{rank}(A)$, is the number of leading 1's in $\text{rref}(A)$.

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

<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 $\text{rank}(A) \leq n$ and $\text{rank}(A) \leq m$.

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

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

(4) If $\text{rank}(A) < m$, 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 $\text{rank}(A) \leq n < m$.

   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.

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

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

Dot product: 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)

Product of a matrix with a vector:

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

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

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 $\mathbb{R}^n$.

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

A vector $\vec{v}$ is a linear combination 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.

Example:

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!

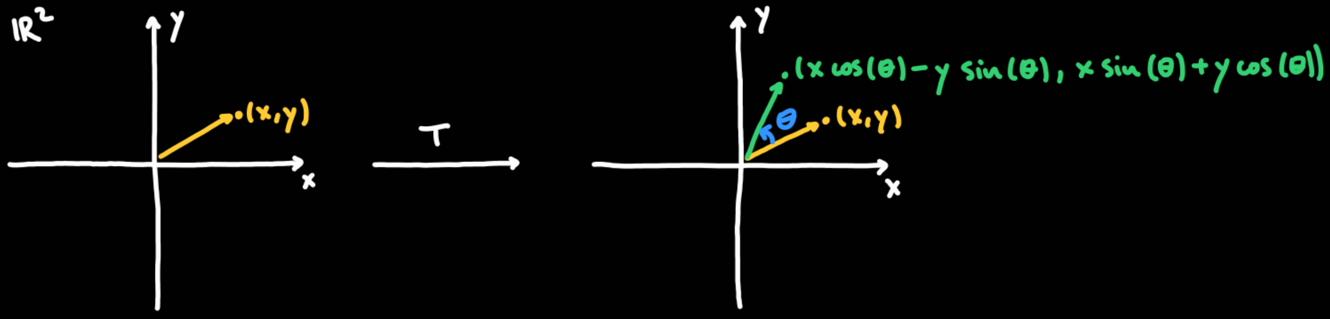
3. $\operatorname{rank}\begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} = 2.$ False!

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.

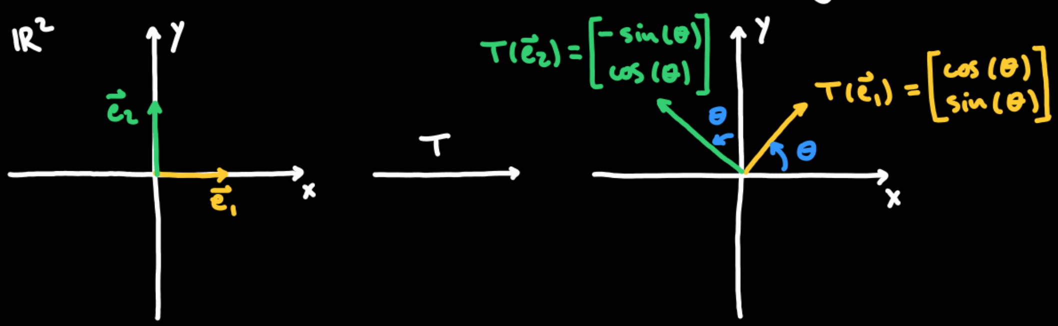


This rotation is a linear transformation because:

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

__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 $\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$.

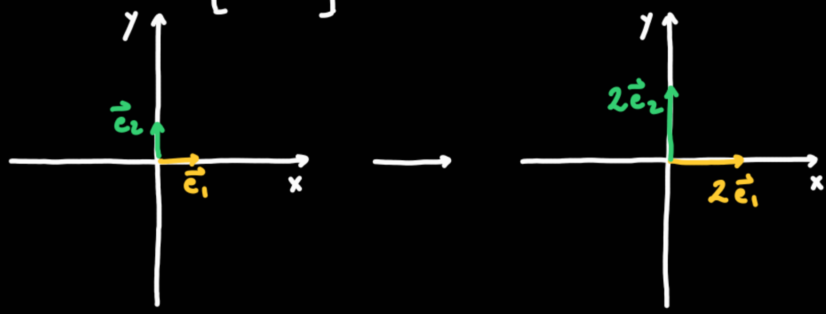
__Theorem:__ A function T from $\mathbb{R}^m$ to $\mathbb{R}^n$ is a linear transformation if and only if:

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

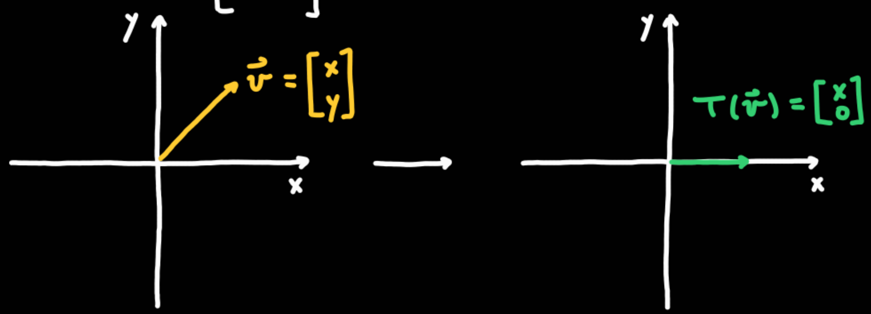
(ii) $T(\lambda \vec{v}) = \lambda T(\vec{v})$ for all $\vec{v}$ in $\mathbb{R}^m$ and λ in $\mathbb{R}$.

__Example:__

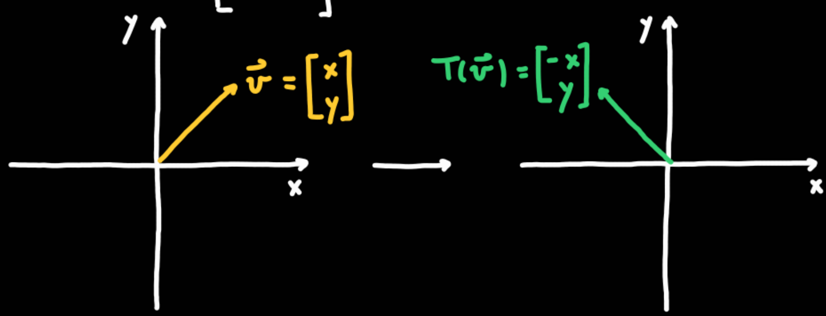
1. The matrix $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ is a dilation by 2 (or scaling).



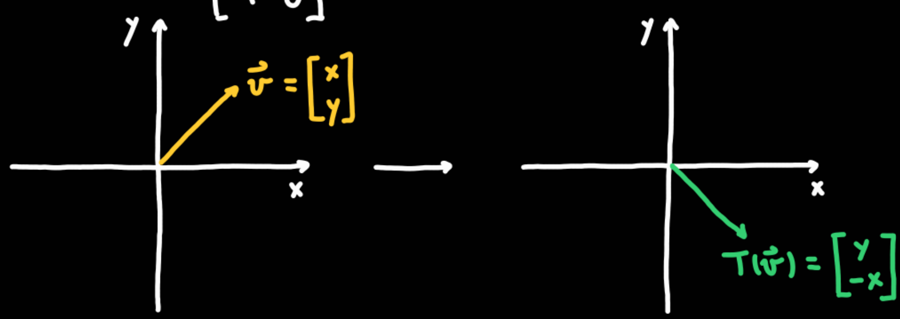
2. The matrix $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ is an orthogonal projection onto the horizontal axis:



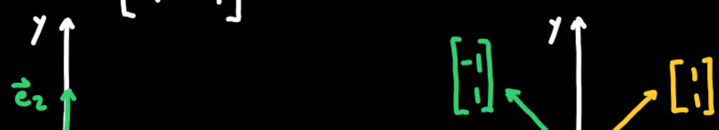
3. The matrix $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ is a reflection about the vertical axis:

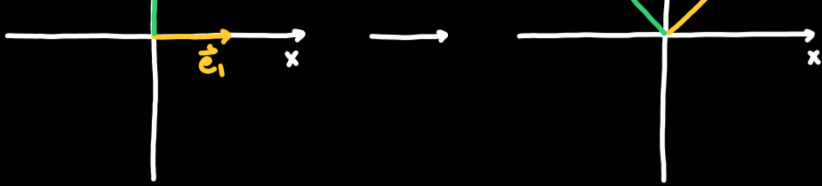


4. The matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is a clockwise rotation of $\frac{\pi}{2}$.



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 $\mathbb{R}$.

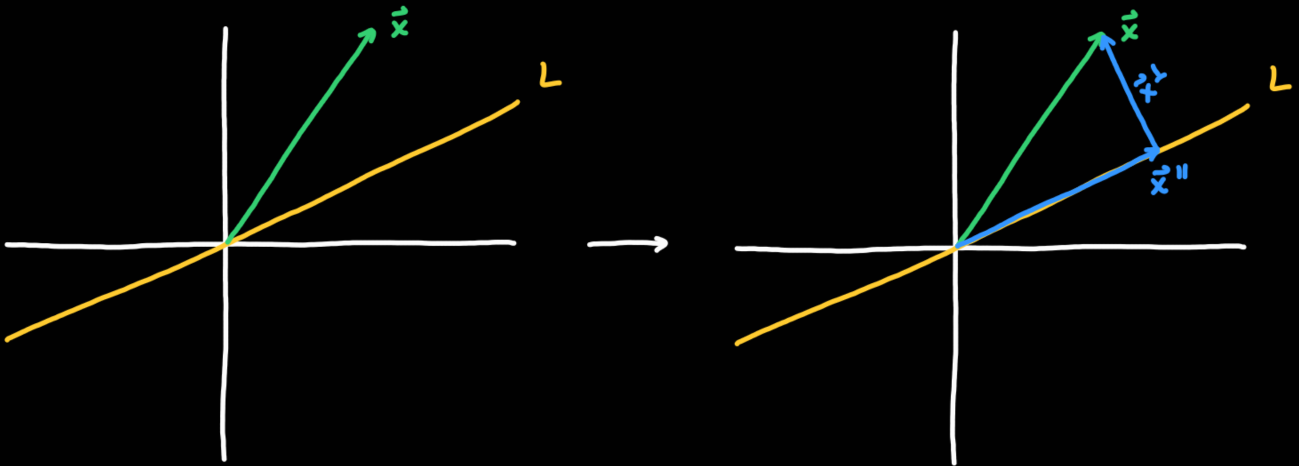
Orthogonal projections.

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

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

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.

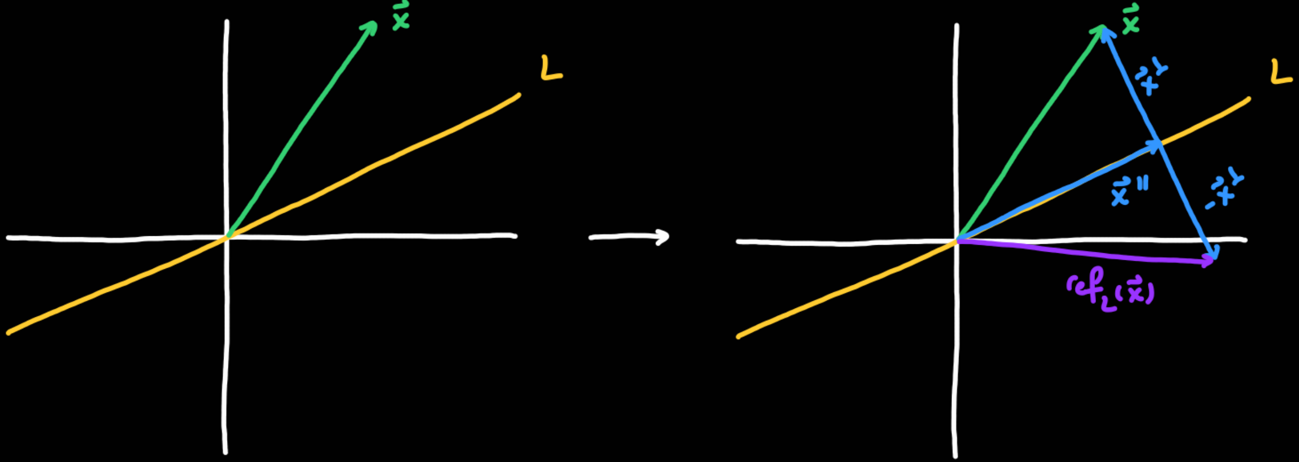


If $\vec{u}$ is non-zero and parallel to L, the associated matrix is $\frac{1}{u_1^2 + u_2^2}\begin{bmatrix} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{bmatrix}$.

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

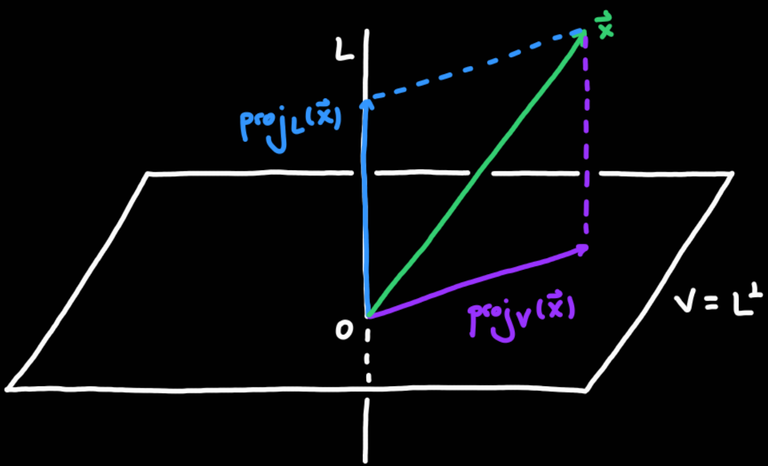
$\text{ref}_L(\vec{x}) = \vec{x}^{\parallel} - \vec{x}^{\perp}$.



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 $\vec{x} = proj_L(\vec{x}) + proj_V(\vec{x})$ where $proj_L(\vec{x})$ is the orthogonal projection of $\vec{x}$ onto L and $proj_V(\vec{x})$ 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) $proj_L(\vec{x}) = (\vec{x}\cdot\vec{u})\vec{u}$

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

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

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

<u>Example:</u>

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

<u>Rotations.</u>

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 $a^2 + b^2 = 1$.

<u>Example:</u> 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:

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

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:

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

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

<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: $(kA)B = A(kB) = k(AB)$.