Review: Linear Algebra

Agenda

- Vector and Matrix
- Vector space
- Vector and matrix operations
- Eigenvalues and Eigenvectors

Motivating example

A system of N=2 linear equations for N=2 unknowns:

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

Has a unique solution,

$$x_1 = \frac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{12} a_{21}}$$
$$x_2 = \frac{b_2 a_{11} - a_{21} b_1}{a_{11} a_{22} - a_{12} a_{21}}$$

Motivating example

Rearrange the numbers in a grid

$$Ax = b$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Apply the determinant rules

$$x_i = \frac{|A_i|}{|A|}$$

$$\det(A) = |A|$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$= a_{11}a_{22} - a_{12}a_{21}$$

$|A_i|$ is formed by replacing the i$^{\text{th}}$ column by the column vector b

Vectors of real numbers R^N

A vector is a set of (ordered) numbers:

$$v = \begin{pmatrix} 1 \\ 6 \\ 3 \\ 5 \end{pmatrix} \in R^4$$

We can switch between "row" and "column" format by transposing:

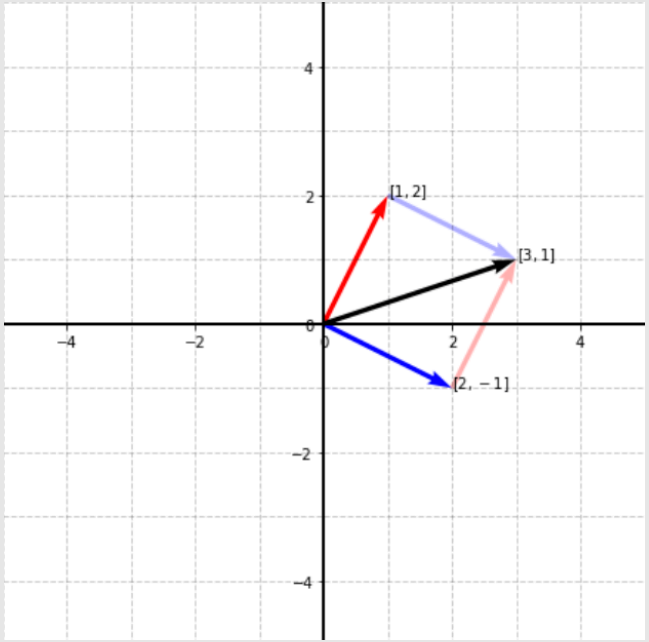
$$v^T = (1, 6, 3, 5)$$

Property 1: vector addition

- Simply add the individual components to obtain result

- E.g.

$$\begin{pmatrix}1\\2\\3\end{pmatrix}+\begin{pmatrix}4\\5\\6\end{pmatrix}=\begin{pmatrix}5\\7\\9\end{pmatrix}$$



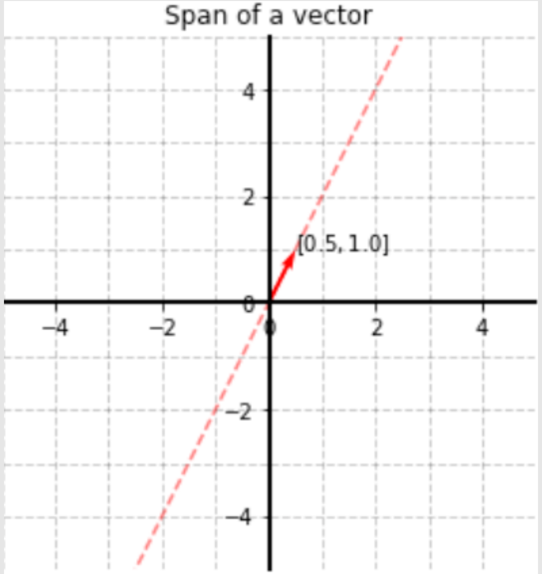
Property 2: vector scaler multiplication

- Multiplying a vector $v \in R^N$ with a scaler $\lambda \in R$, elongates it:

$$v \to \lambda v$$

- As we vary λ, we obtain a line called the span of the vector

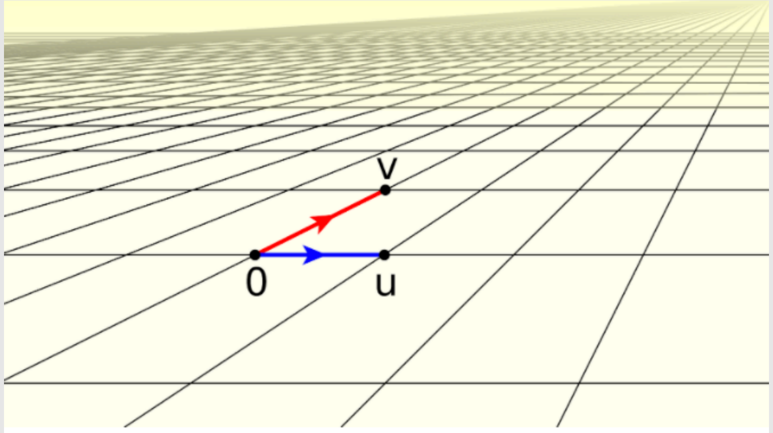
$$\text{span}(v) = \{\lambda v |\, \lambda \in R\}$$



Span of a set of vectors

- If S = $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_k}\}$ is a set of vectors in a vector space V, then the span of S is the set of all linear combinations of the vectors in S.

$$\text{span}(S) = \{c_1\boldsymbol{v_1} + c_2\,\boldsymbol{v_2} + \cdots + c_k\boldsymbol{v_k}, c_1, c_2, \dots, c_k \in R\}$$



Matrix: a generalization of vectors

Matrix is a $M \times N$ object of real numbers

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Matrix transpose (Flip rows and columns)

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}^T = \begin{bmatrix} a_{1,1} & a_{2,1} & a_{3,1} \\ a_{1,2} & a_{2,2} & a_{3,2} \\ a_{1,3} & a_{2,3} & a_{3,3} \end{bmatrix}$$

Matrix properties

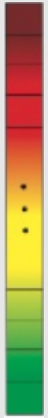
Matrix addition

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{bmatrix} = \begin{bmatrix} a_{1,1}+b_{1,1} & a_{2,1}+b_{2,1} & a_{3,1}+b_{3,1} \\ a_{1,2}+b_{1,2} & a_{2,2}+b_{2,2} & a_{3,2}+b_{3,2} \\ a_{1,3}+b_{1,3} & a_{2,3}+b_{2,3} & a_{3,3}+b_{3,3} \end{bmatrix}^T$$

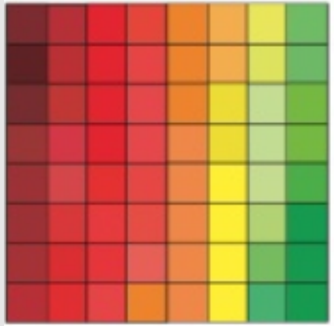
Scalar multiplication

$$\lambda \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} = \begin{bmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \lambda a_{1,3} \\ \lambda a_{2,1} & \lambda a_{2,2} & \lambda a_{2,3} \\ \lambda a_{3,1} & \lambda a_{3,2} & \lambda a_{3,3} \end{bmatrix}$$

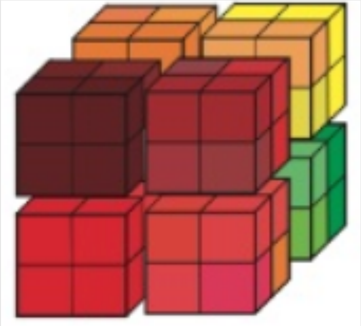
Tensors: a generalization of matrix



Vector



Matrix
$X \in R^{8\times 8}$



Tensor
$X \in R^{4\times 4\times 4}$

Vector spaces

Vector space: definition

Definition: A *vector space* is a set V on which two operations + and · are defined, called *vector addition* and *scalar multiplication*.

The operation + (vector addition) must satisfy the following conditions:

Closure: If $\mathbf{u}$ and $\mathbf{v}$ are any vectors in V, then the sum $\mathbf{u} + \mathbf{v}$ belongs to V.

(1) *Commutative law:* For all vectors $\mathbf{u}$ and $\mathbf{v}$ in V, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

(2) *Associative law:* For all vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in V, $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$

(3) *Additive identity:* The set V contains an *additive identity* element, denoted by $\mathbf{0}$, such that for any vector $\mathbf{v}$ in V, $\mathbf{0} + \mathbf{v} = \mathbf{v}$ and $\mathbf{v} + \mathbf{0} = \mathbf{v}$.

(4) *Additive inverses:* For each vector $\mathbf{v}$ in V, the equations $\mathbf{v} + \mathbf{x} = \mathbf{0}$ and $\mathbf{x} + \mathbf{v} = \mathbf{0}$ have a solution $\mathbf{x}$ in V, called an *additive inverse* of $\mathbf{v}$, and denoted by - $\mathbf{v}$.

The operation · (scalar multiplication) is defined between real numbers (or scalars) and vectors, and must satisfy the following conditions:

Closure: If $\mathbf{v}$ in any vector in V, and c is any real number, then the product $c \cdot \mathbf{v}$ belongs to V.

(5) *Distributive law:* For all real numbers c and all vectors $\mathbf{u}$, $\mathbf{v}$ in V, $c \cdot (\mathbf{u} + \mathbf{v}) = c \cdot \mathbf{u} + c \cdot \mathbf{v}$

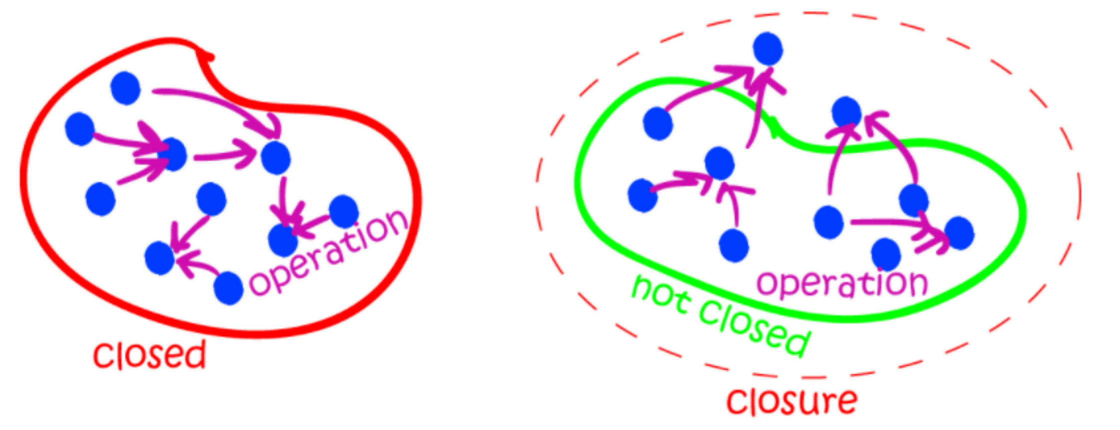
(6) *Distributive law*: For all real numbers c, d and all vectors $\mathbf{v}$ in V, $(c+d) \cdot \mathbf{v} = c \cdot \mathbf{v} + d \cdot \mathbf{v}$

(7) *Associative law*: For all real numbers c,d and all vectors $\mathbf{v}$ in V, $c \cdot (d \cdot \mathbf{v}) = (cd) \cdot \mathbf{v}$

(8) *Unitary law*: For all vectors $\mathbf{v}$ in V, $1 \cdot \mathbf{v} = \mathbf{v}$

Closure

What is the set of all things that can result from the proposed operation

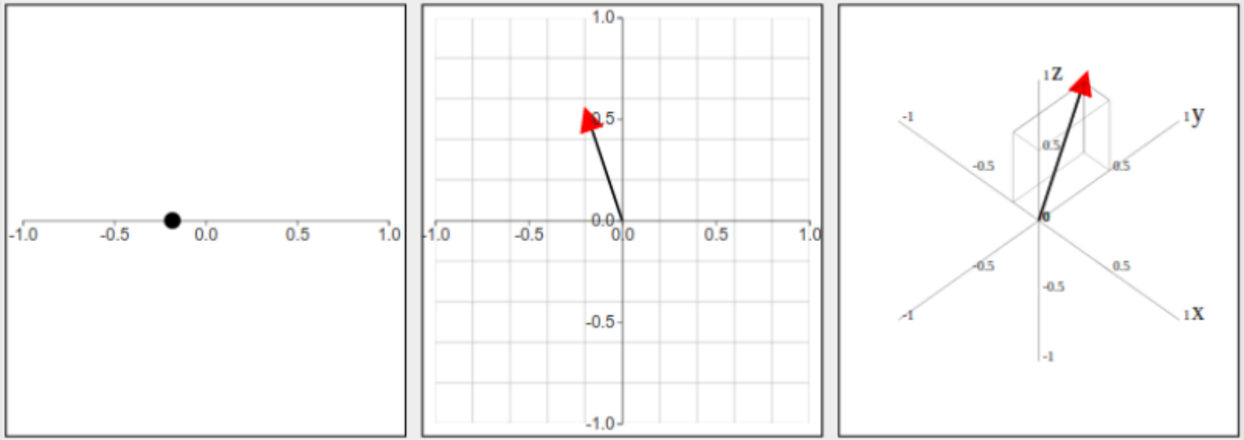


Real numbers are closed under addition and subtraction

Positive numbers are not closed under subtraction

Vector space

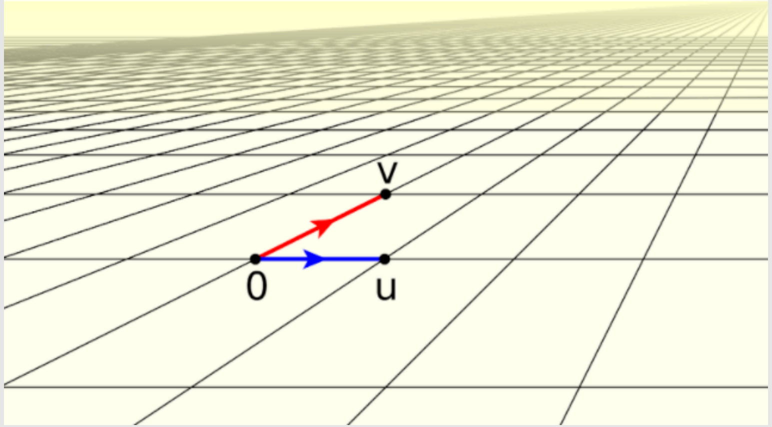
A vector space V is a set that is closed under finite vector addition and scalar multiplication



Euclidean n-space R^n is called a real vector space

N vectors can generate a vector space

For a vector space V, $\mathcal{A} = \{x_1, x_2, \ldots, x_n\}$ is a generating set if $V = span(\mathcal{A})$



$$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$$

A basis is a **generating** set that is **minimal**

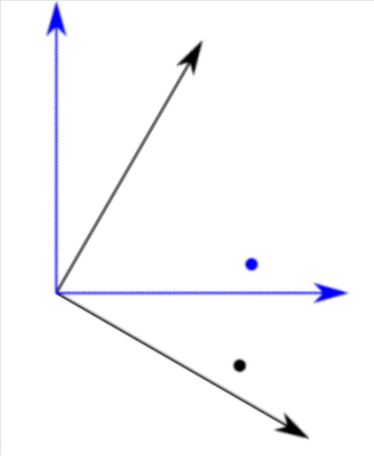
Vectors correspond to coordinates w.r.t. a basis

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

$$\boldsymbol{x} = x_1 \boldsymbol{i}_1 + x_2 \boldsymbol{i}_2 + \cdots + x_n \boldsymbol{i}_n$$

Change of basis

Sometimes it's useful to change the basis



New coordinates will be

$$x' = M^{-1}x \text{ where } M = [j_1, j_2, \dots, j_n]$$

Vector and Matrix Operations

Vector-Vector Products

For any two vectors:

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_n \end{bmatrix} \quad \text{u}= \begin{bmatrix} u_1 \\ u_2 \\ \cdots \\ u_n \end{bmatrix}$$

The **dot product** is $u \cdot v = \sum_{i=1}^{n} u_i v_i$

Sometime we write the dot product as $u^T v$ or $v^T u$

Two vectors are orthogonal if their dot product is 0

An inner/dot product allows us to define a norm

To measure the size of a vector $\|\cdot\|: V \to R$

$$\|x\|_1 = \sum_{j=1}^{N} |x_j|$$

$$\|x\|_2 = \sqrt{\sum_{j=1}^{N} x_j^2} = \sqrt{x^T x}$$

Or more generally $$\|x\|_p = \left(\sum_{j=1}^{N} |x_j|^p\right)^{1/p}$$

We can calculate the distance between two vectors as

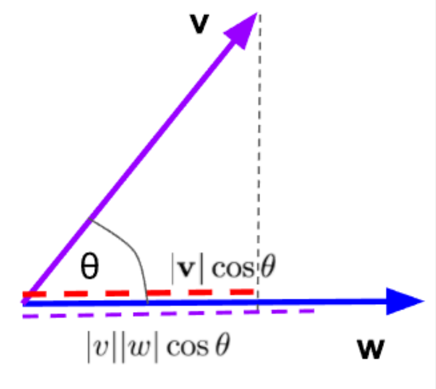
$$d(u, v) = \|u - v\|$$

Dot Product and Geometry

Once you have the concept of a norm, you can derive

$$v \cdot w = \|v\| \|w\| cos\theta$$

$$\|v\| cos\theta = v \cdot \frac{w}{\|w\|}$$



The dot product is fundamentally projecting vector v on a rescaled unit vector in the direction of w

Matrix Vector Product

Given a matrix $A \in R^{m \times n}$ and a vector $x \in R^n$

$$Ax = \begin{bmatrix} — & a_1^T & — \\ — & a_2^T & — \\ & \vdots & \\ — & a_m^T & — \end{bmatrix} x = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_m^T x \end{bmatrix}$$

Matrix Product

$$c_{11}=a_{11}b_{11}+a_{12}b_{21}+a_{13}b_{31}+a_{14}b_{41}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}$$

2×4 4×3 2×3

Linear Independent and Rank

A set of vectors $\{x_1, x_2, \dots x_n\}$ are said to be (linearly) independent if no vector can be represented as a linear combination of the remaining vectors.

A set of vectors are said to be (linearly) dependent if

$$x_n = \sum_{i=1}^{n-1} a_i x_i \quad \text{for some } a_1, \dots a_{n-1}$$

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \; x_2 = \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix} x_3 = \begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}$$

Linearly dependent $x_3 = -2x_1 + x_2$

Linear Independent and Rank

For a matrix matrix $A \in R^{m \times n}$, the column rank is the size of the largest subset of columns that constitute a linear independent set.
The row rank is the largest number of rows that constitute a linearly independent set.
For any matrix $A \in R^{m \times n}$, the column rank is always equal to the row rank.

For $A \in R^{m \times n}$, $Rank(A) \leq \min(m, n)$
If $Rank(A) = \min(m, n)$, then A is said to be Full Rank

Inverse of a square matrix

For a square matrix $A \in R^{n \times n}$, the inverse A^{-1} is the unique matrix such that

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

Only square matrices can have the inverse.

If A^{-1} exists, we say that A is invertible or non-singular

In order for a square matrix to have an inverse A^{-1}, then A must be of full rank

Vector transformation and eigenvectors

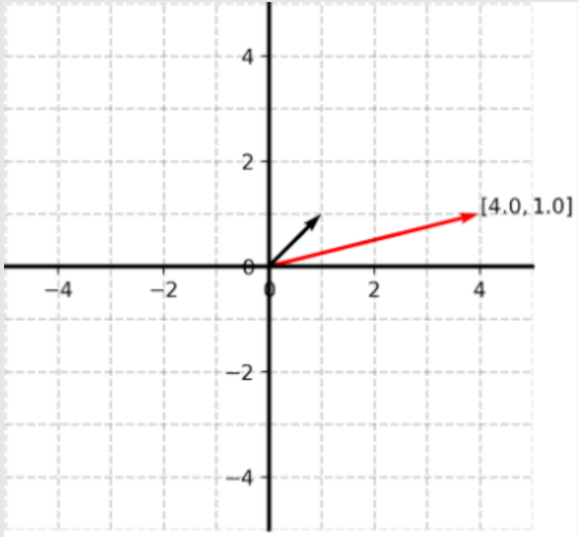
Transformation of vectors

functions that map a vector space V into a vector space W, denoted by T: V → W

If **v** is in V and **w** is in W such that $T(\boldsymbol{v}) = \boldsymbol{w}$, then **w** is the image of **v** under T.

E.g. $T: v \rightarrow \begin{bmatrix} 3v_1 + v_2 \\ v_1 \end{bmatrix}$

Then: $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$



Linear Transformation Rules:

$$T(\boldsymbol{v} + \boldsymbol{w}) = T(\boldsymbol{v}) + T(\boldsymbol{w})$$
$$T(\lambda\, \boldsymbol{x}) = \lambda T(\boldsymbol{x})$$

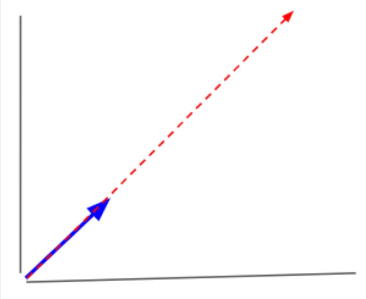
Matrix: representation of a linear transformation

In R^N, a linear transformation of a vector can be represented by a matrix:

$$Tv$$

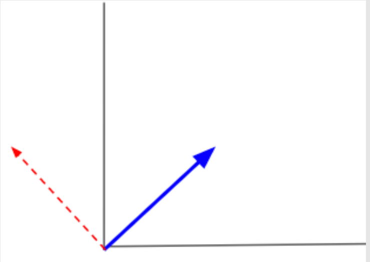
Stretching

$$\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$



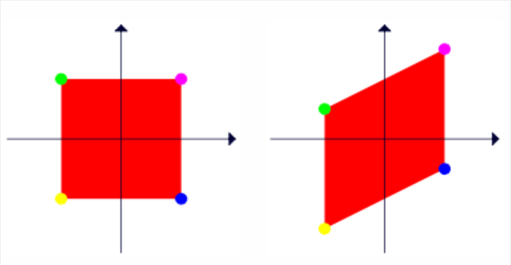
Rotation

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

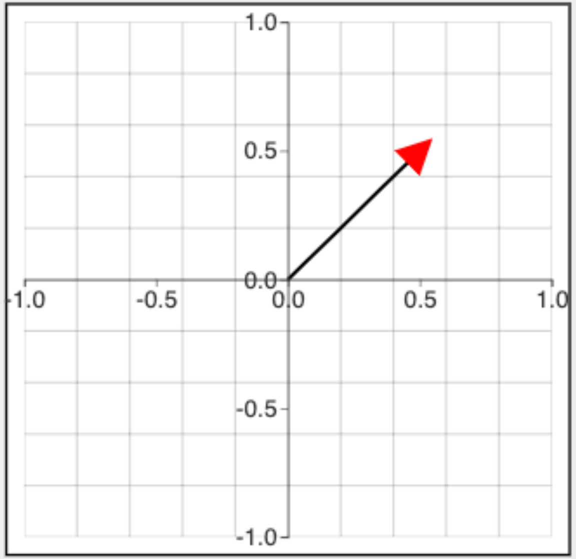


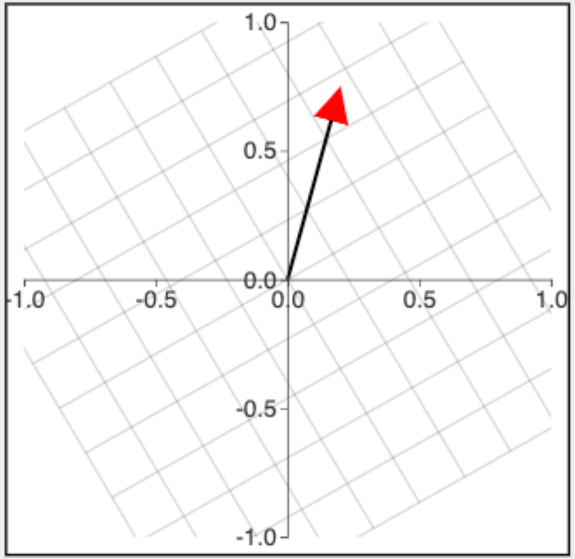
Shearing

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$



Example



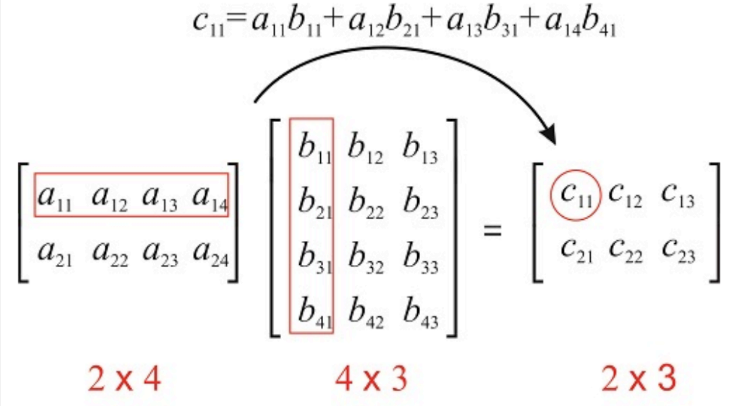


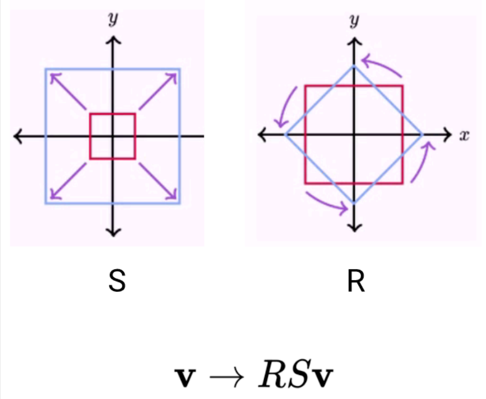
The rotation transform changes the directions but does not affect the size of the vectors:

$$\begin{bmatrix} x \\ y \end{bmatrix} \leftarrow \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

A random matrix can be constructed as a sequence of rotations, scaling, projections and shearing

Basic matrix operations





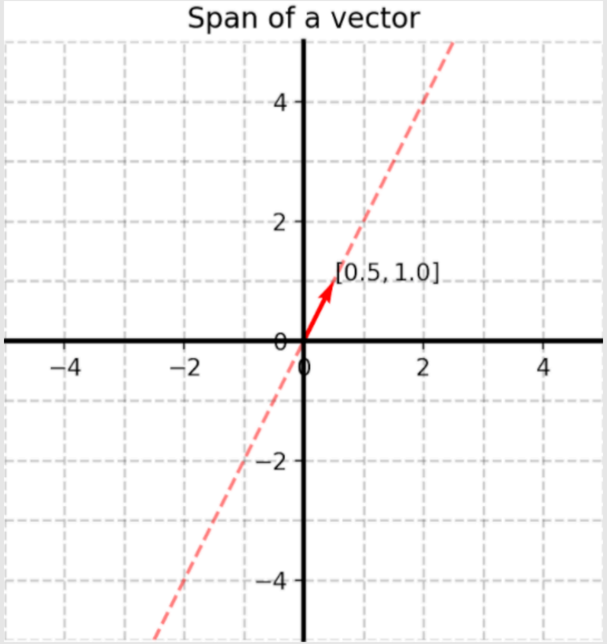
Multiplying matrices = composing transformations one after another

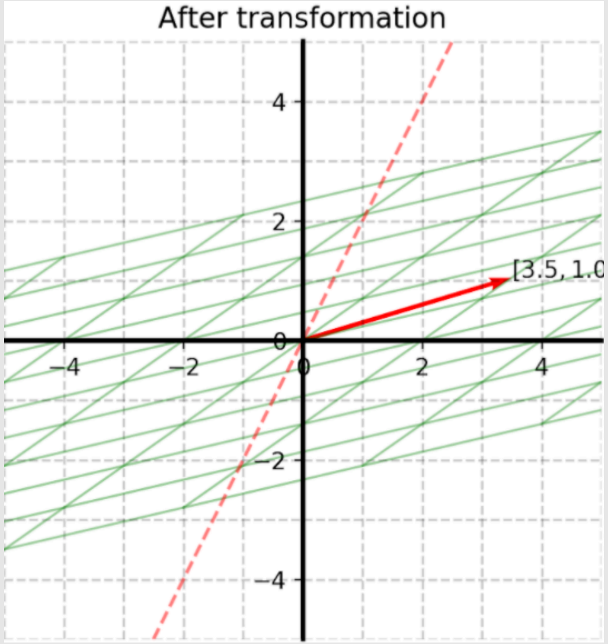
Caveat: Linear Algebra

Linear algebra means that after transformation:

All straight lines must remain straight lines without getting curved. All grid lines have to stay parallel and evenly spaced. The origin has to stay put.

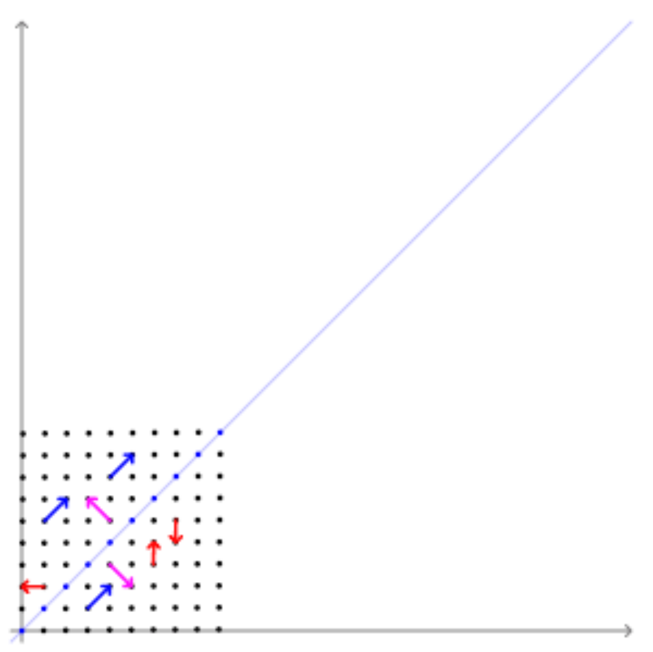
Transforming space: span(v)





The vector will usually not lie on its span after transformation

Transforming space: eigenvectors



Transformation matrix

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Eigen vectors:

$$(1, 1)^T$$
$$(1, -1)^T$$

Eigenvalues: 3, 1

In some cases the vector remains on its own span

Eigenvectors

Eigenvectors are vectors that stay on their span after transformation T is applied to them

$$Tu = \lambda u$$

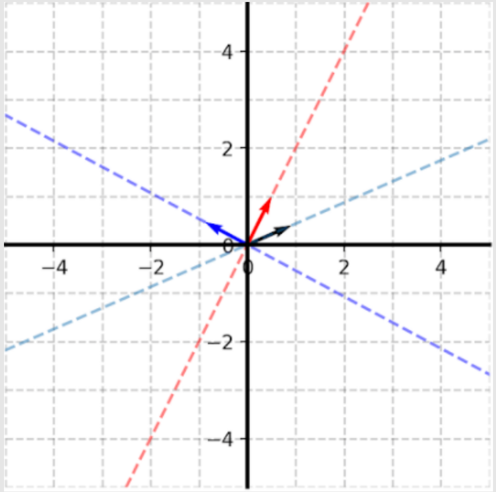
T: original transformation matrix

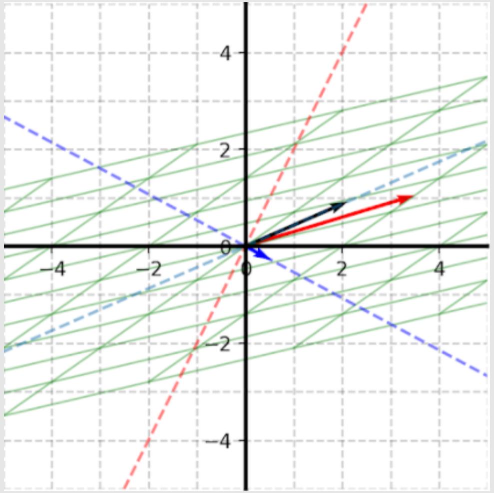
u: eigenvectors (not unique)

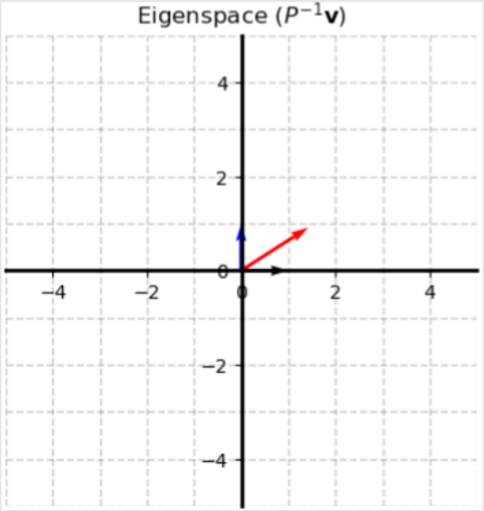
λ: eigenvalues

Eigenvectors: what's the deal?

Eigenvectors can create their own "eigenspace" if you can use them as a basis







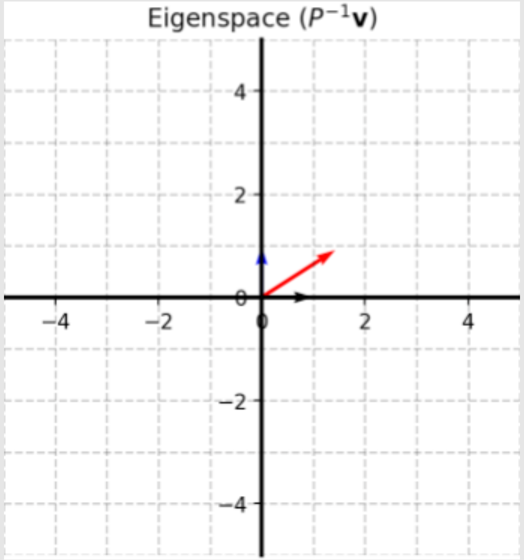
Transformation by matrix
T in original space

Note change of basis is easily obtained by P^{-1} as: $P = [e_1, e_2]$

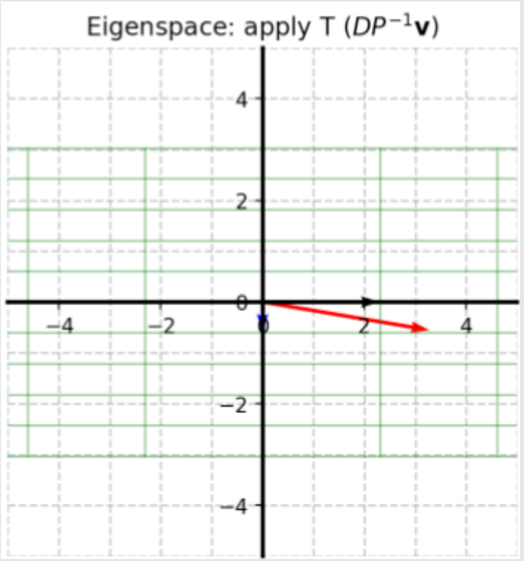
Eigenvectors: what's the deal

In eigenspace, the transformation is elegantly clean:

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$



Transformation by matrix T in original space



Transformation by matrix D in Eigenspace

Why is this so?

We know that

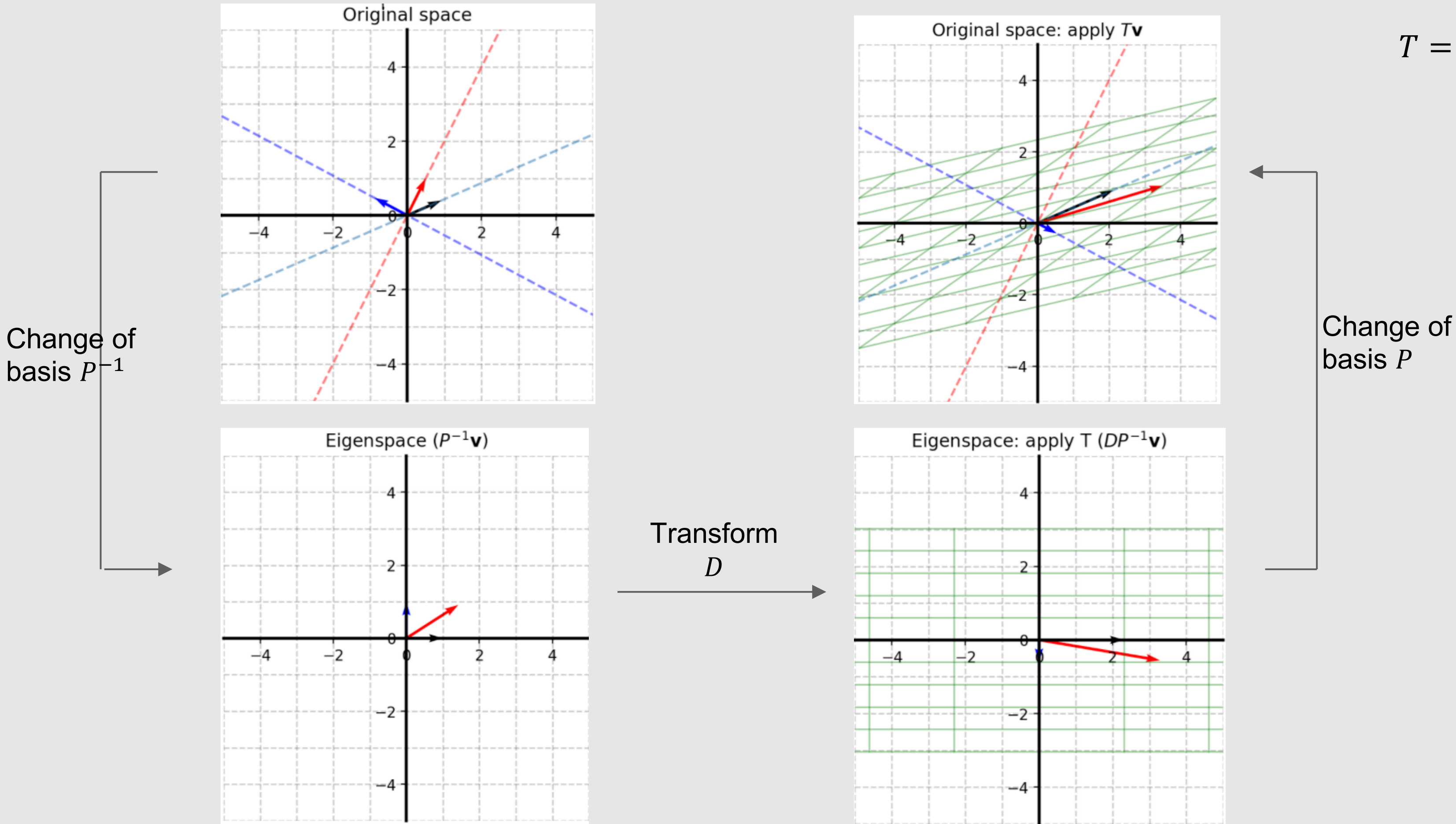
$$\begin{cases} Te_1 = \lambda_1 e_1 \\ Te_2 = \lambda_2 e_2 \\ \qquad \dots \end{cases}$$

It can be rewritten as

$$TP = PD$$
$$T = PDP^{-1}$$

Where $P = [e_1, e_2, \dots, e_N]$ $D = \begin{bmatrix} \lambda_1 & & \\ & \cdots & \\ & & \lambda_N \end{bmatrix}$

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


Original space
Original space: apply $T\mathbf{v}$
Change of basis P^{-1}
Change of basis P
Eigenspace ($P^{-1}\mathbf{v}$)
Transform D
Eigenspace: apply T ($DP^{-1}\mathbf{v}$)