Machine Learning CMPT 410/726

Mo Chen
SFU School of Computing Science
2025-09-05

Linear Algebra and Calculus Review

Vectors

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4$$

Addition

$$\vec{x} + \vec{y} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

Scaling

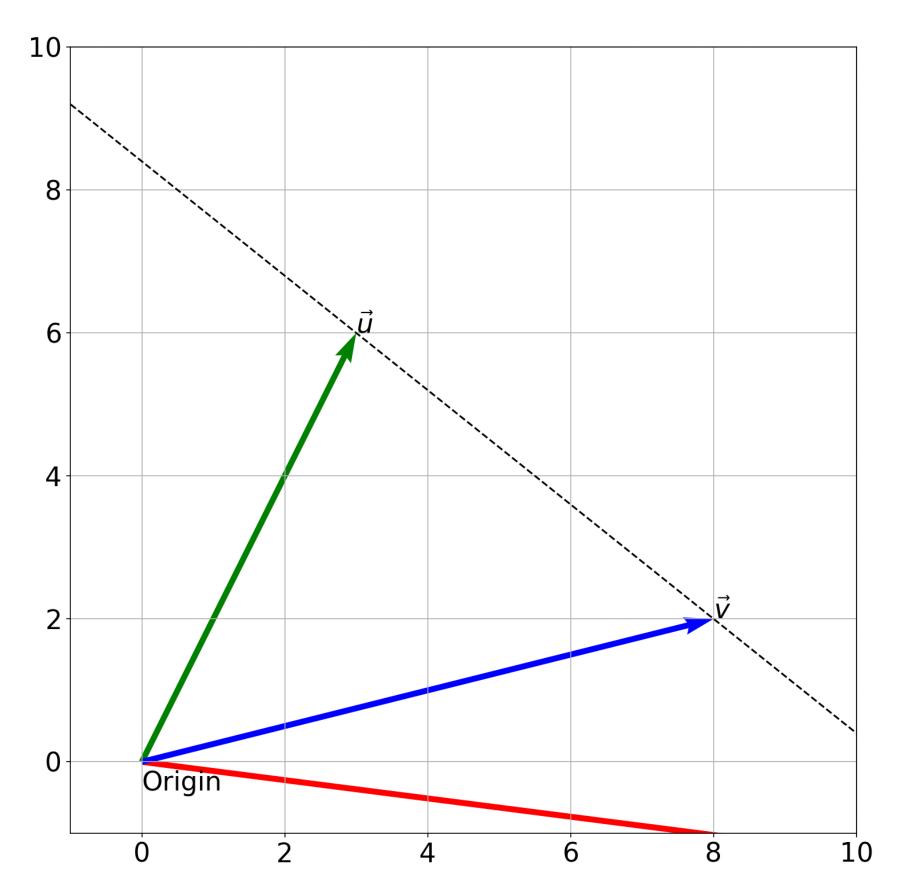
$$c\vec{x} = \begin{pmatrix} cx_1 \\ cx_2 \\ cx_3 \\ cx_4 \end{pmatrix}$$

$$\vec{x} = \alpha_1 \vec{v}_1 + \dots + \alpha_n \vec{v}_n = \sum_{i=1}^N \alpha_i \vec{v}_i$$

Special Case:

$$\sum_{i=1}^{N} \alpha_i = 1$$

Called "Affine Combination"



$$\alpha = -0.9$$

$$\beta = 1.9$$
subject to $\alpha + \beta = 1$

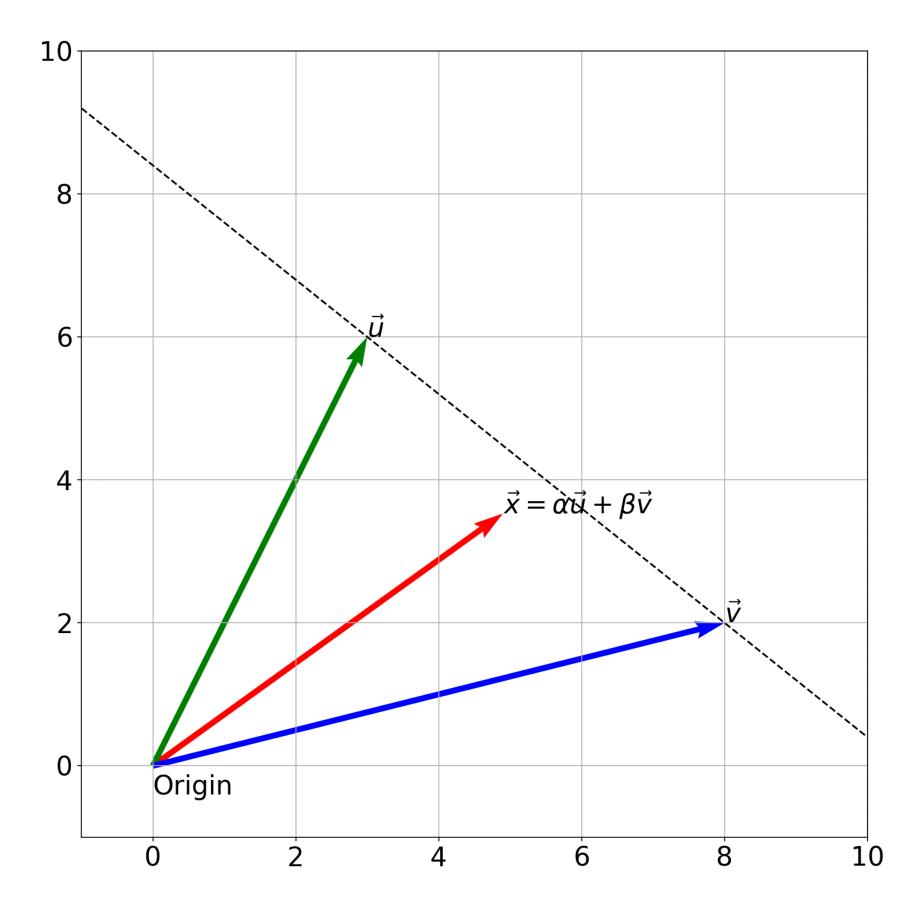
$$\vec{x} = \alpha \vec{u} + \beta \vec{v}$$

$$\vec{x} = \alpha \vec{u} + \beta \vec{v}$$

$$\vec{x} = \alpha_1 \vec{v}_1 + \dots + \alpha_n \vec{v}_n = \sum_{i=1}^N \alpha_i \vec{v}_i$$

Special Case:

$$\sum_{i=1}^{N} \alpha_i < 1$$



 $\alpha = 0.44$

 $\beta = 0.45$

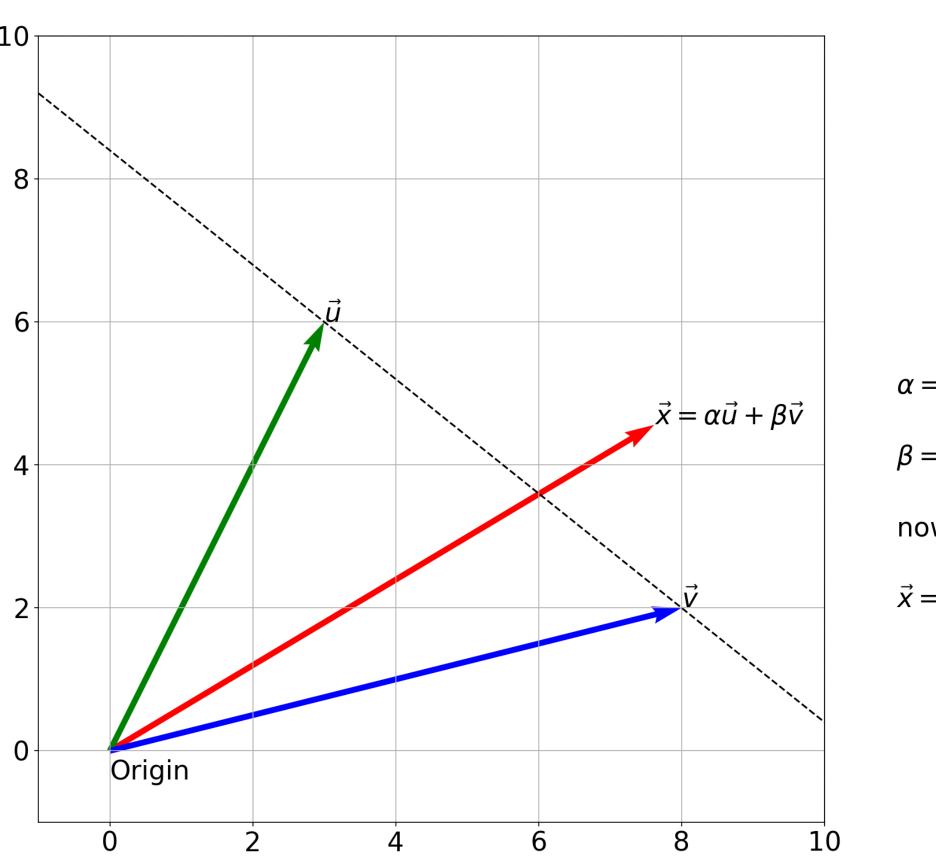
now $\alpha + \beta < 1$

 $\vec{x} = \alpha \vec{u} + \beta \vec{v}$

$$\vec{x} = \alpha_1 \vec{v}_1 + \dots + \alpha_n \vec{v}_n = \sum_{i=1}^N \alpha_i \vec{v}_i$$

Special Case:

$$\sum_{i=1}^{N} \alpha_i > 1$$



 $\alpha = 0.51$

 $\beta = 0.76$

now $\alpha + \beta > 1$

 $\vec{x} = \alpha \vec{u} + \beta \vec{v}$

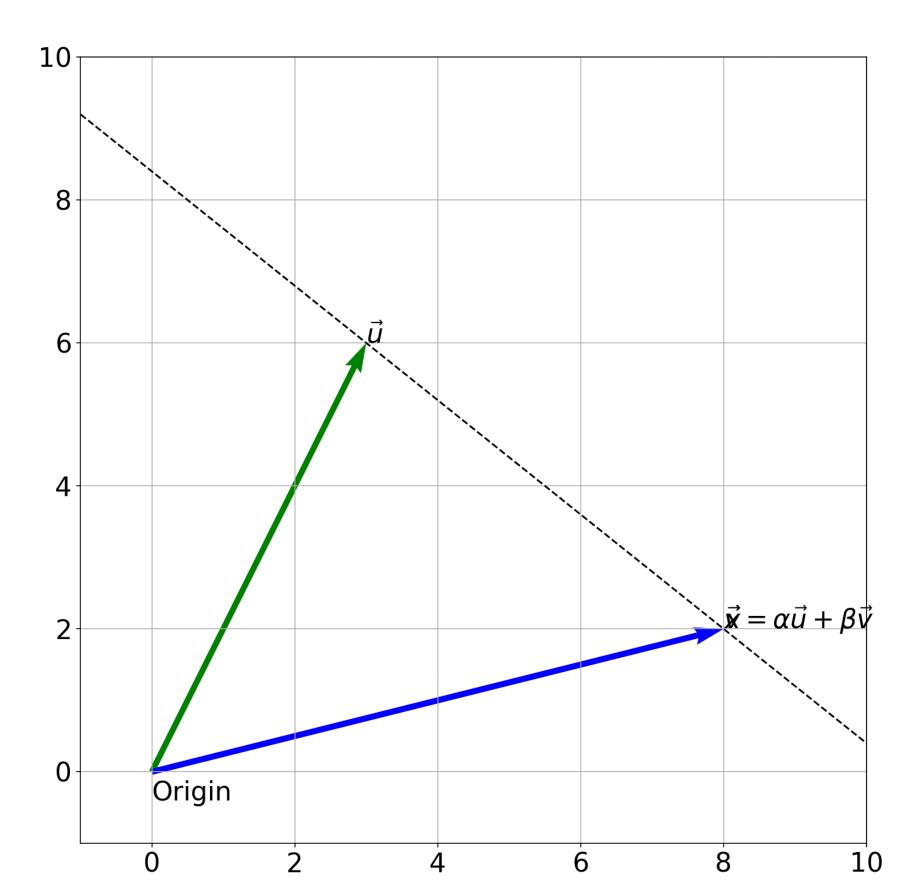
$$\vec{x} = \alpha_1 \vec{v}_1 + \dots + \alpha_n \vec{v}_n = \sum_{i=1}^N \alpha_i \vec{v}_i$$

Special Case:

$$\sum_{i=1}^{N} \alpha_i = 1$$

$$\alpha_i \geq 0 \ \forall i$$

Called "Convex Combination"



$$\alpha = -0.0$$

$$\beta = 1.0$$
subject to $\alpha + \beta = 1$

$$\vec{x} = \alpha \vec{u} + \beta \vec{v}$$

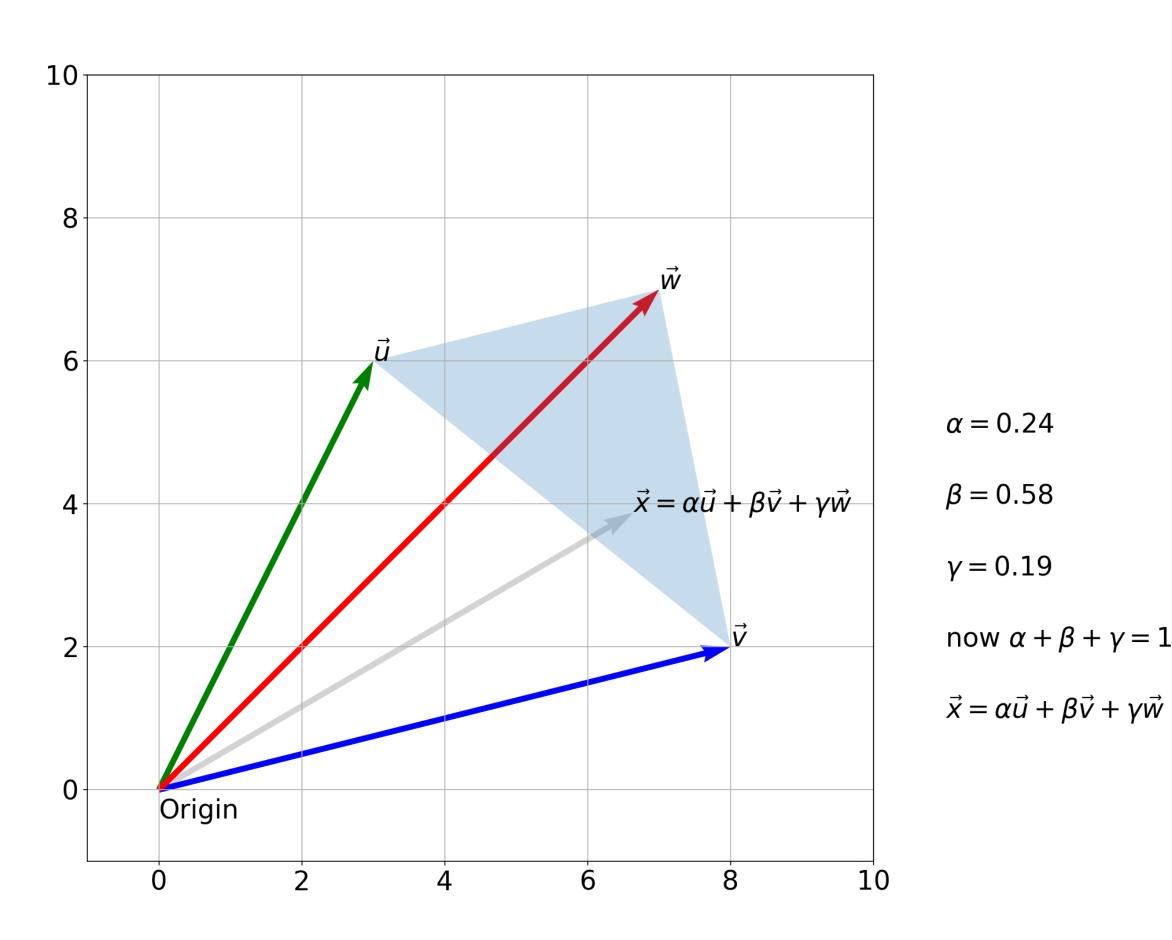
$$\vec{x} = \alpha_1 \vec{v}_1 + \dots + \alpha_n \vec{v}_n = \sum_{i=1}^N \alpha_i \vec{v}_i$$

Special Case:

$$\sum_{i=1}^{N} \alpha_i = 1$$

$$\alpha_i \geq 0 \ \forall i$$

Called "Convex Combination"



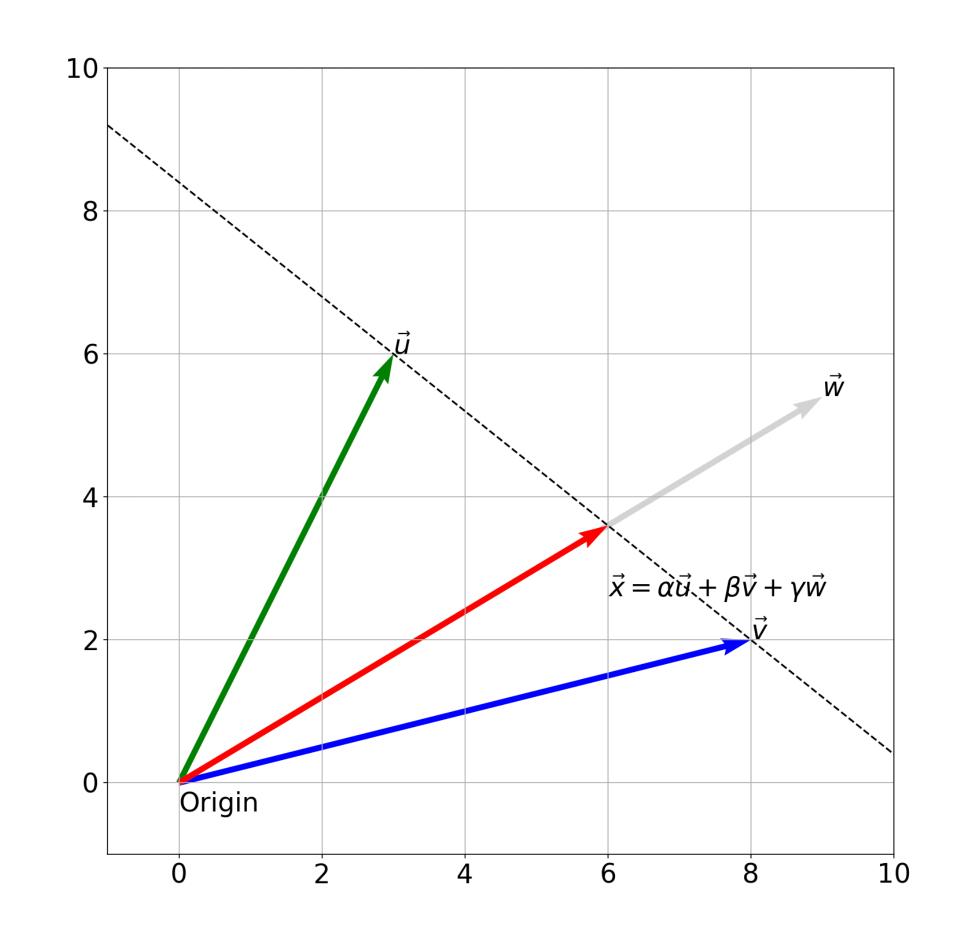
Span and Linear Independence

$$\operatorname{span}\{\vec{v}_1,\ldots,\vec{v}_n\} = \left\{\sum_{i=1}^N \alpha_i \vec{v}_i \mid \alpha_1,\ldots,\alpha_n \in \mathbb{R}\right\}$$

A set of vectors is *linearly* independent if no vector is in the span of the other vectors.

$$\sum_{i=1}^{N} \alpha_i \vec{v}_i = \vec{0} \Longrightarrow \alpha_1, \dots, \alpha_n = 0$$

Otherwise, they are *linearly* dependent.



 \vec{w} can be represented by combination of \vec{u} and \vec{v} starting from $\alpha + \beta = 1$ and adjust proportionally $\alpha = 0.4$ $\beta = 0.6$ $\vec{x} = \alpha \vec{u} + \beta \vec{v}$

Inner Products*

$$\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^{N} x_i y_i$$

For the special case of real vectors:

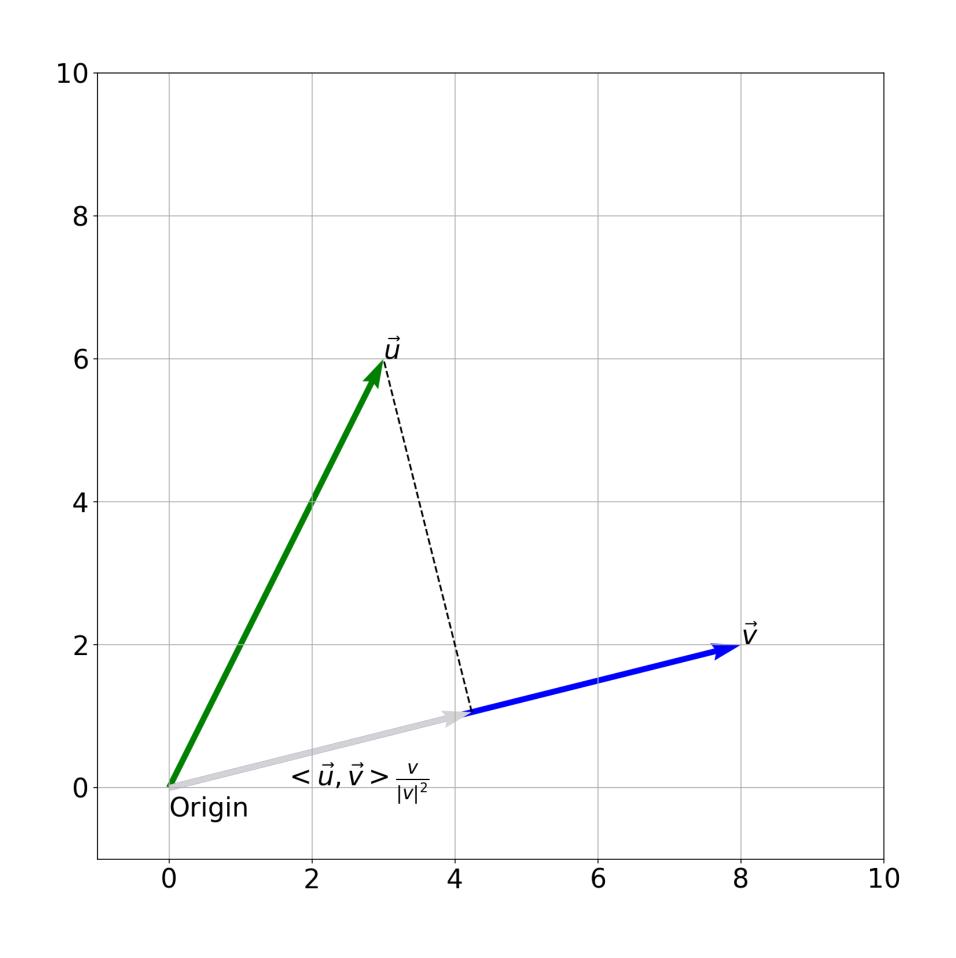
Symmetry: $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$

Bilinearity:

$$\langle \alpha \vec{x} + \beta \vec{y}, \vec{z} \rangle = \alpha \langle \vec{x}, \vec{z} \rangle + \beta \langle \vec{y}, \vec{z} \rangle$$

$$\langle \vec{x}, \alpha \vec{y} + \beta \vec{z} \rangle = \alpha \langle \vec{x}, \vec{y} \rangle + \beta \langle \vec{x}, \vec{z} \rangle$$

Nonnegativity: $\langle \vec{x}, \vec{x} \rangle \ge 0 \ \forall \vec{x}$



^{*}Strictly speaking, this is the special case of standard inner products

Norms

Euclidean norm:

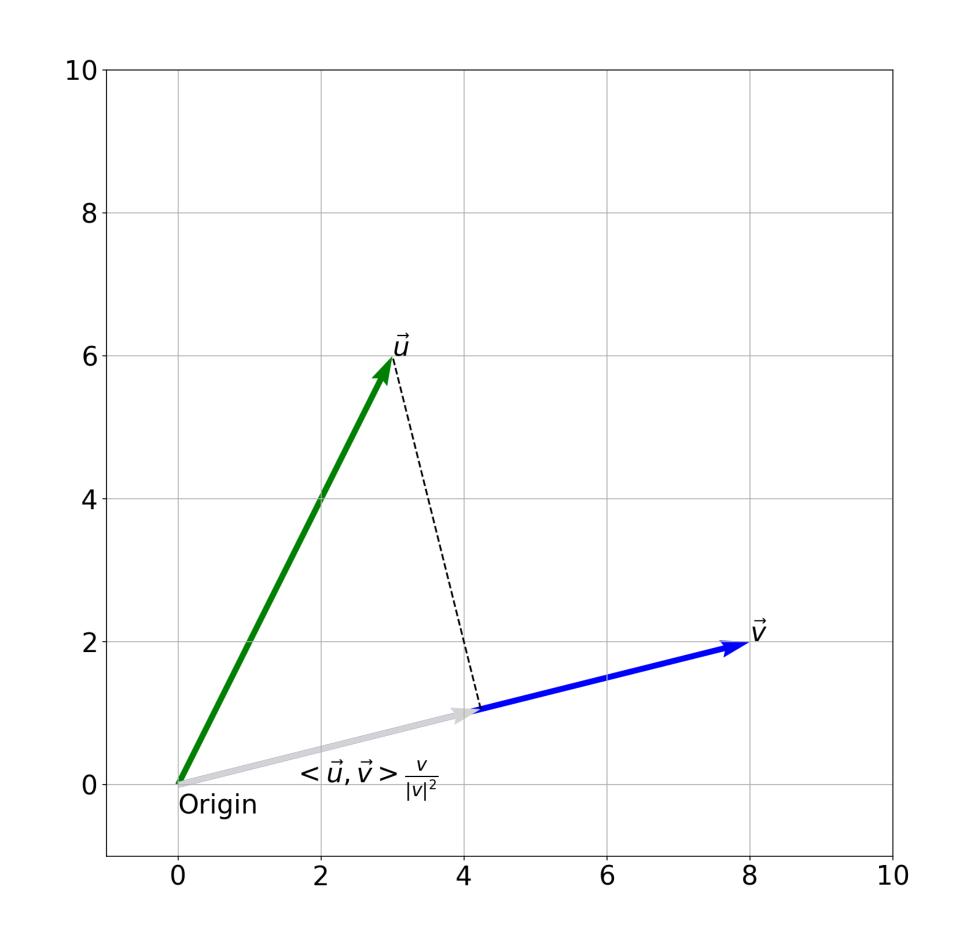
$$\|\vec{x}\|_2 = \sqrt{\langle \vec{x}, \vec{x} \rangle} = \sqrt{\sum_{i=1}^n x_i^2}$$
$$\|\vec{x}\|_2 \ge 0 \ \forall \vec{x}$$

Unit vector: vector with norm 1,

often referred to as "unit norm" Normalizing a vector \vec{x} : $\frac{\vec{x}}{\|\vec{x}\|_2}$

Projecting a vector \vec{x} onto \vec{y} :

$$\left\langle \vec{x}, \frac{\vec{y}}{\|\vec{y}\|_2} \right\rangle \frac{\vec{y}}{\|\vec{y}\|_2}$$



Sign of $\langle \vec{x}, \vec{y} \rangle$ is positive if angle is less than 90 degrees, and negative otherwise. 11

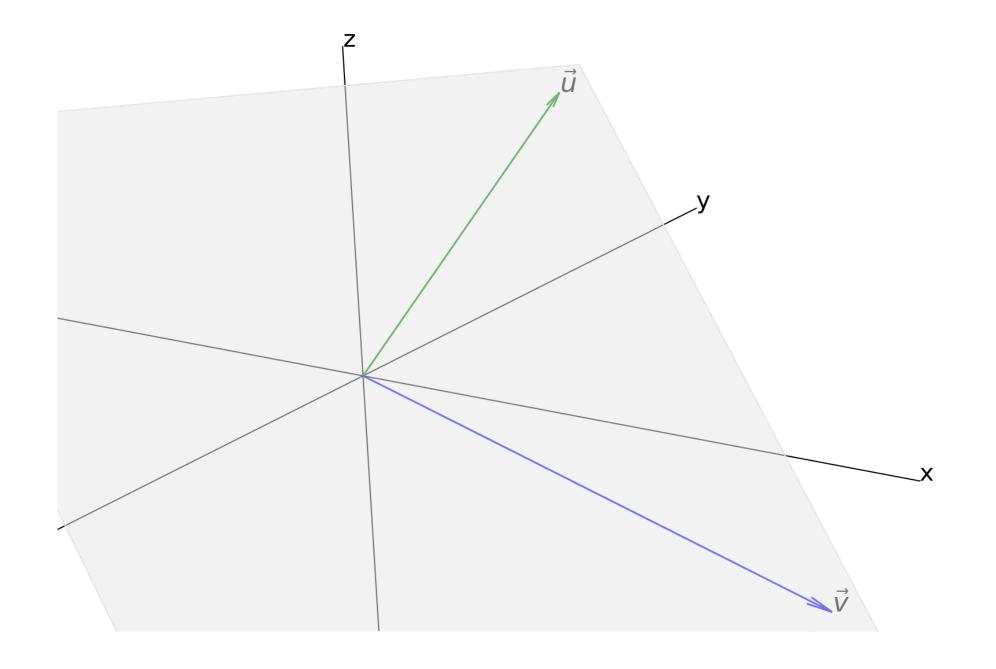
Linear Subspace

A lower-dimensional slice of the vector space that passes through the origin.

Examples:

A plane through the origin in 3D Euclidean space

A line through the origin in 3D Euclidean space

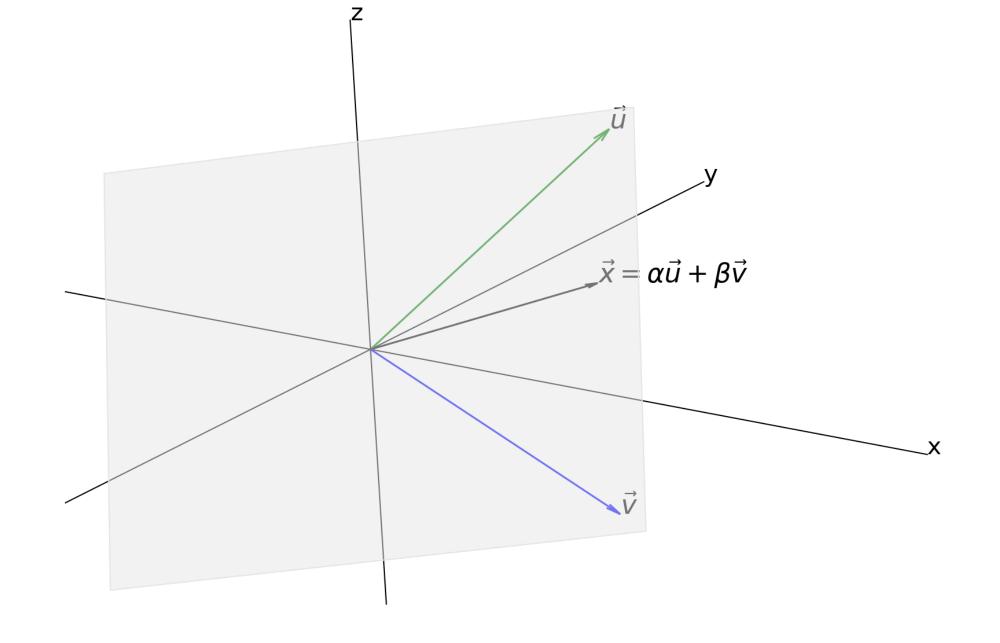


Basis

A linearly independent set of vectors that spans the full (sub)space.

Any vector in the (sub)space can be written as a linear combination of the basis vectors.

 $\overrightarrow{u}, \overrightarrow{v}$ are a basis for the subspace



$$\alpha = 0.57$$

$$\beta = 0.36$$

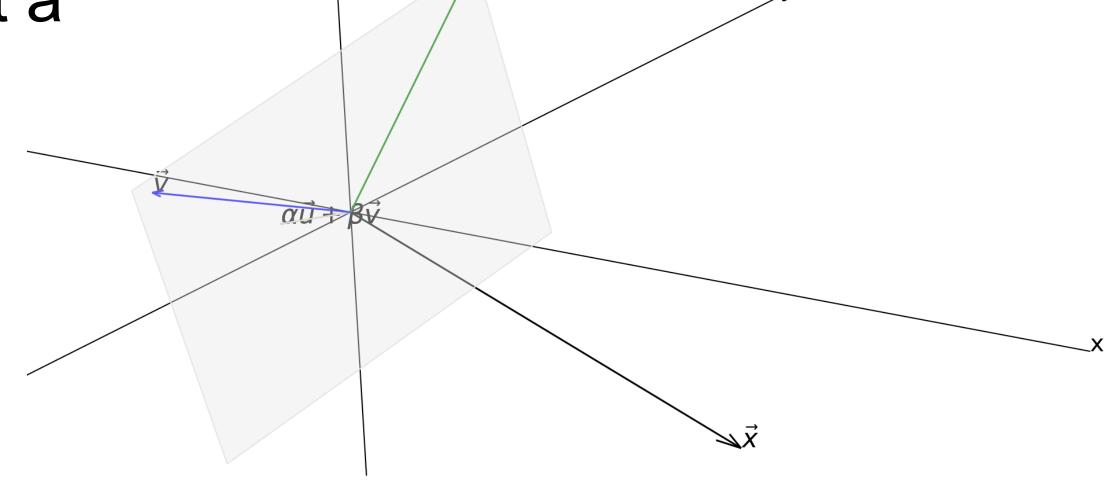
$$\vec{x} = \alpha \vec{u} + \beta \vec{v}$$

Basis

A linearly independent set of vectors that spans the full (sub)space.

Any vector in the (sub)space can be written as a linear combination of the basis vectors.

 $\overrightarrow{u}, \overrightarrow{v}$ are not a basis for \mathbb{R}^3



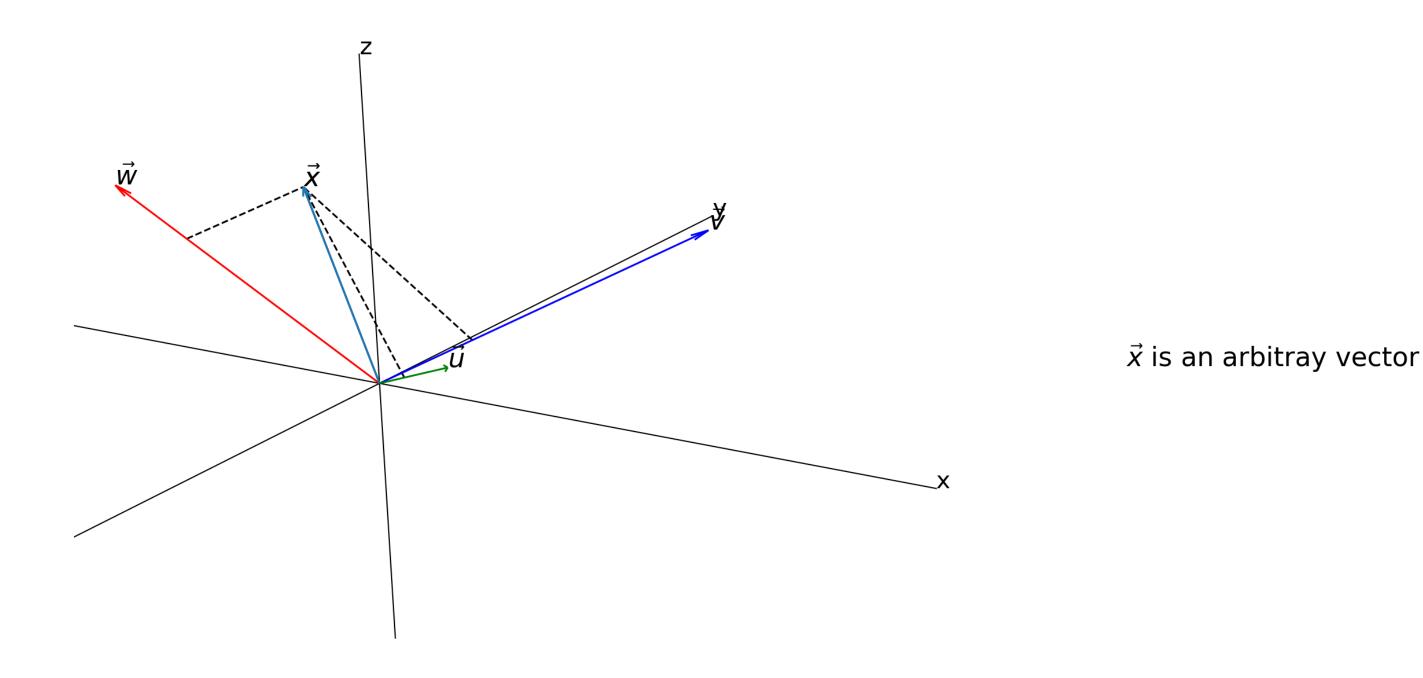
 $\not\exists \alpha, \beta, s. t.$

 $\vec{x} = \alpha \vec{u} + \beta \vec{v}$

Orthogonal Basis

Two vectors \vec{x} , \vec{y} are orthogonal if and only if $\langle \vec{x}, \vec{y} \rangle = 0$.

An orthogonal basis is a basis whose vectors are orthogonal to one another.



Orthogonal Basis

We like an orthogonal basis because it is easy to compute the coordinates with respect to the basis.

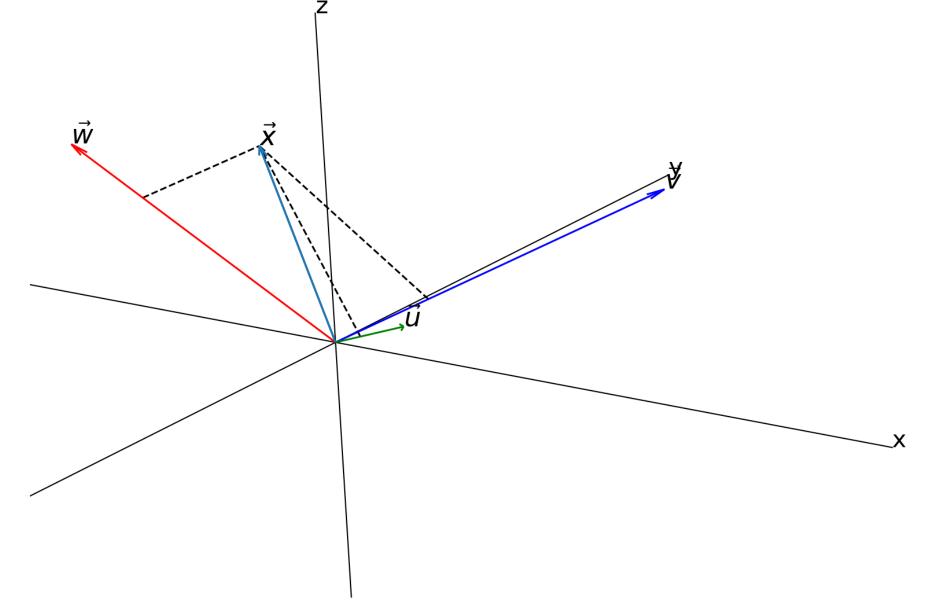
$$\langle \vec{x}, \vec{v}_j \rangle = \left\langle \sum_{i=1}^{N} \alpha_i \vec{v}_i, \vec{v}_j \right\rangle = \sum_{i=1}^{n} \alpha_i \langle \vec{v}_i, \vec{v}_j \rangle = \alpha_j \langle \vec{v}_j, \vec{v}_j \rangle = \alpha_j \|\vec{v}_j\|_2^2 \Rightarrow \alpha_j = \frac{\langle \vec{x}, \vec{v}_j \rangle}{\|\vec{v}_j\|_2^2}$$

$$\vec{x} \text{ is an arbitray vector}$$

Orthonormal Basis

An orthonormal basis is a special case of an orthogonal basis with unit vectors.

$$\langle \vec{x}, \vec{v}_j \rangle = \left\langle \sum_{i=1}^n \alpha_i \vec{v}_i, \vec{v}_j \right\rangle = \sum_{i=1}^n \alpha_i \langle \vec{v}_i, \vec{v}_j \rangle = \alpha_j \langle \vec{v}_j, \vec{v}_j \rangle = \alpha_j \|\vec{v}_j\|_2^2 \Longrightarrow \alpha_j = \frac{\langle \vec{x}, \vec{v}_j \rangle}{\|\vec{v}_j\|_2^2}$$



$$\|\vec{v}_j\|_2 = 1 \Longrightarrow \alpha_j = \langle \vec{x}, \vec{v}_j \rangle$$

 \vec{x} is an arbitray vector

Finding the coordinates gets even easier!

Standard Basis

The standard basis is an orthonormal basis with the following vectors:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Usually denoted as:

$$\vec{e}_1, \vec{e}_2, \vec{e}_3, \dots, \vec{e}_{n-1}, \vec{e}_n$$

Known as the standard basis vectors.

Matrices

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \\ a_{4,1} & a_{4,2} & a_{4,3} \end{pmatrix} \in \mathbb{R}^{4 \times 3}$$

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \\ a_{4,1} & a_{4,2} & a_{4,3} \end{pmatrix} \in \mathbb{R}^{4 \times 3}$$

Short-hand:

$$A = \left(a_{ij}\right) \in \mathbb{R}^{m \times n}$$

$$A + B = (a_{ij} + b_{ij})$$
$$cA = (ca_{ij})$$

Multiplying a matrix with a vector:

$$\vec{y} = A\vec{x}$$

$$\begin{pmatrix} ?? \\ ?? \\ ?? \\ ?? \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 \\ 1 & 1 & -1 \\ 2 & 0 & 4 \\ -2 & 0 & -4 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$$

Multiplying a matrix with a matrix:

$$AX = A \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ x_{21} & \cdots & x_{2n} \\ x_{31} & \cdots & x_{3n} \end{pmatrix}$$
$$= A(\vec{x}_{\cdot 1} & \cdots & \vec{x}_{\cdot n})$$
$$= (A\vec{x}_{\cdot 1} & \cdots & A\vec{x}_{\cdot n})$$

$$(AB)C = A(BC)$$
, but $AB \neq BA$

Matrices

Transpose: A^{T}

If
$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \\ a_{4,1} & a_{4,2} & a_{4,3} \end{pmatrix}$$

then
$$A^{\mathsf{T}} = \begin{pmatrix} a_{1,1} & a_{2,1} & a_{3,1} & a_{4,1} \\ a_{1,2} & a_{2,2} & a_{3,2} & a_{4,2} \\ a_{1,3} & a_{2,3} & a_{3,3} & a_{4,3} \end{pmatrix}$$

Will treat vectors as column vectors:

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Vector transpose denotes a row vector:

$$\vec{x}^{\mathsf{T}} = (x_1 \quad \cdots \quad x_n)$$

Outer product:

$$\vec{x}^{\top}\vec{y} = (x_1 \quad \cdots \quad x_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \sum_{i=1}^n x_i y_i = \langle \vec{x}, \vec{y} \rangle$$

$$\vec{x}\vec{y}^{\top} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} (y_1 & \cdots & y_n)$$

$$= \begin{pmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \vdots & \vdots & \ddots & \vdots \\ x_ny_1 & x_ny_2 & \cdots & x_ny_n \end{pmatrix}$$

$$(A^{\top})^{\top} = A$$

$$(A + B)^{\top} = A^{\top} + A^{\top$$

Properties of Transpose:

$$(A^{\mathsf{T}})^{\mathsf{T}} = A$$
$$(A + B)^{\mathsf{T}} = A^{\mathsf{T}} + B^{\mathsf{T}}$$
$$(cA)^{\mathsf{T}} = cA^{\mathsf{T}}$$
$$(AB)^{\mathsf{T}} = B^{\mathsf{T}}A^{\mathsf{T}}$$

Matrices

Rank: Number of linearly independent columns, or equivalently the number of linearly independent rows, denoted as rank(A)

$$A = \begin{pmatrix} -1 & 2 \\ 0 & 0 \\ 2 & -4 \end{pmatrix}$$
 rank $(A) = 1$ $B = \begin{pmatrix} -1 & 2 \\ 0 & 1 \\ 2 & -4 \end{pmatrix}$ rank $(B) = 2$ Full-rank

$$B = \begin{pmatrix} -1 & 2 \\ 0 & 1 \\ 2 & -4 \end{pmatrix} \quad \begin{array}{l} \operatorname{rank}(B) = 2 \\ \operatorname{Full-rank} \end{array}$$

Identity Matrix:

$$I = egin{pmatrix} 1 & 0 & 0 & \cdots & 0 \ 0 & 1 & 0 & \cdots & 0 \ 0 & 0 & 1 & \cdots & 0 \ dots & dots & dots & dots & dots \ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

Property: $I\vec{x} = \vec{x}$

Inverse:

A matrix A^{-1} such that $A^{-1}A = I$

Or equivalently,

A matrix A^{-1} such that $AA^{-1} = I$

Not all matrices are invertible! Only those that are square and full-rank are.

Interpreting Matrices

• Each matrix represents a linear transformation, that is, a linear function that maps a vector to a vector.

$$A = (\vec{a}_{\cdot 1} \cdots \vec{a}_{\cdot n})$$

$$A\vec{x} = (\vec{a}_{\cdot 1} \cdots \vec{a}_{\cdot n}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 \vec{a}_{\cdot 1} + \dots + x_n \vec{a}_{\cdot n}$$

- This is a linear combination of $\vec{a}_{.1}, \dots, \vec{a}_{.n}$, where the coefficients are given by \vec{x}
 - Before: $\vec{x} = x_1 \vec{e}_1 + \dots + x_n \vec{e}_n$
 - After: $A\vec{x} = x_1\vec{a}_{\cdot 1} + \dots + x_n\vec{a}_{\cdot n}$

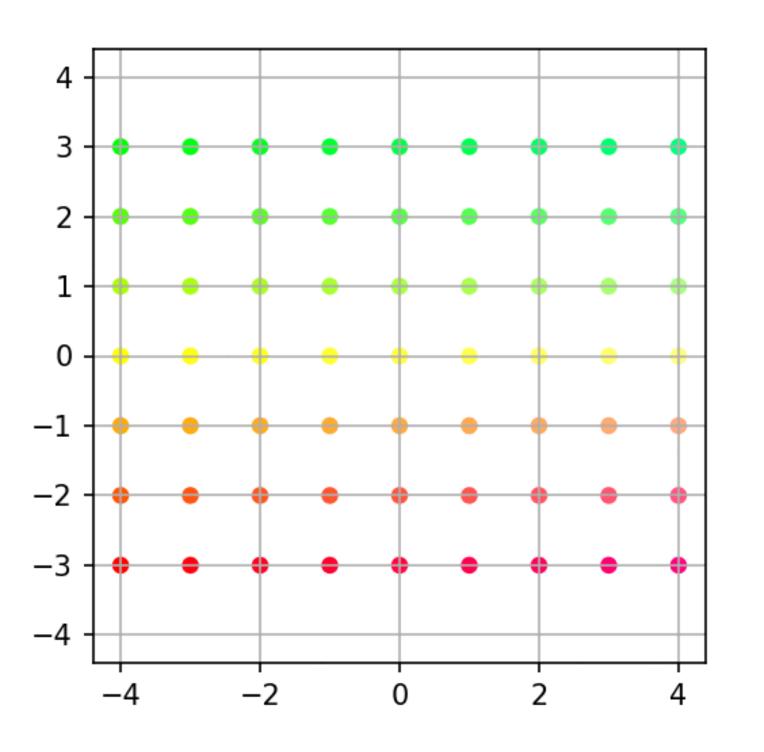
Essentially replacing the original basis vectors with the columns of *A*

$$A = \begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{pmatrix}$$

$$A\vec{x} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 \\ a_{2,2}x_2 \end{pmatrix}$$

In the case of diagonal matrices, rank is the number of non-zero entries.

Full-rank:
$$A = \begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{pmatrix}$$



0.9 0.0

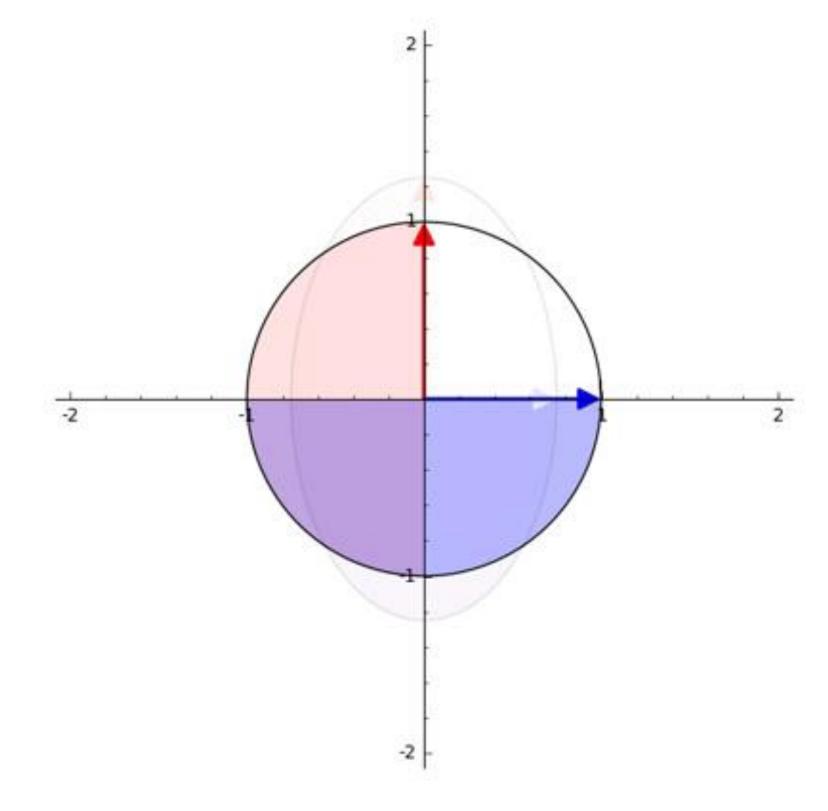
0.0 0.5

$$A = \begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{pmatrix}$$

$$A\vec{x} = A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 \\ a_{2,2}x_2 \end{pmatrix}$$

In the case of diagonal matrices, rank is the number of non-zero entries.

Full-rank:
$$A = \begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{pmatrix}$$



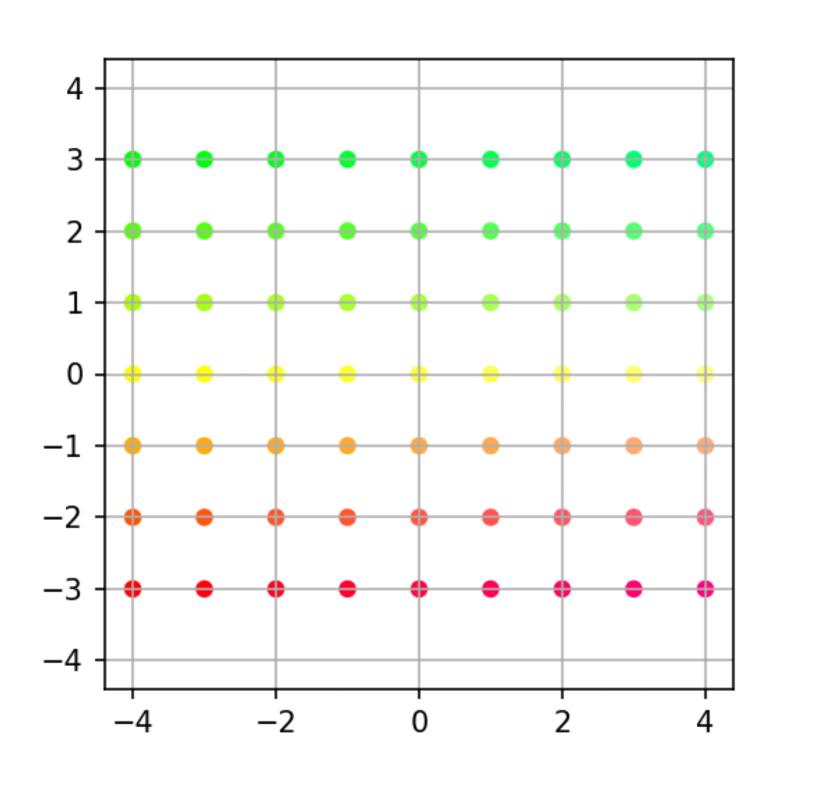
Credit: Ryan Holbrook

$$A = \begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{pmatrix}$$

$$A\vec{x} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 \\ a_{2,2}x_2 \end{pmatrix}$$

In the case of diagonal matrices, rank is the number of non-zero entries.

Rank-deficient:
$$A = \begin{pmatrix} 0 & 0 \\ 0 & a_{2,2} \end{pmatrix}$$



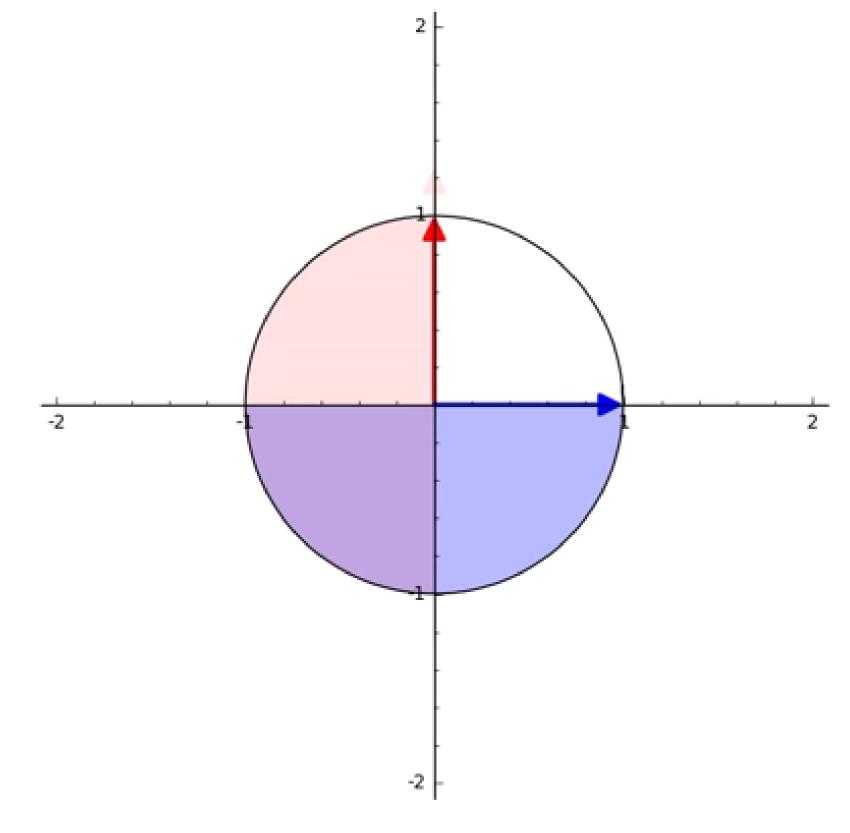
0.0 0.0

0.0 0.5

$$A = \begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{pmatrix}$$
$$A\vec{x} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} a_{1,1}x_1 \\ a_{2,2}x_2 \end{pmatrix}$$

In the case of diagonal matrices, rank is the number of non-zero entries.

Rank-deficient:
$$A = \begin{pmatrix} 0 & 0 \\ 0 & a_{2,2} \end{pmatrix}$$



Credit: Ryan Holbrook

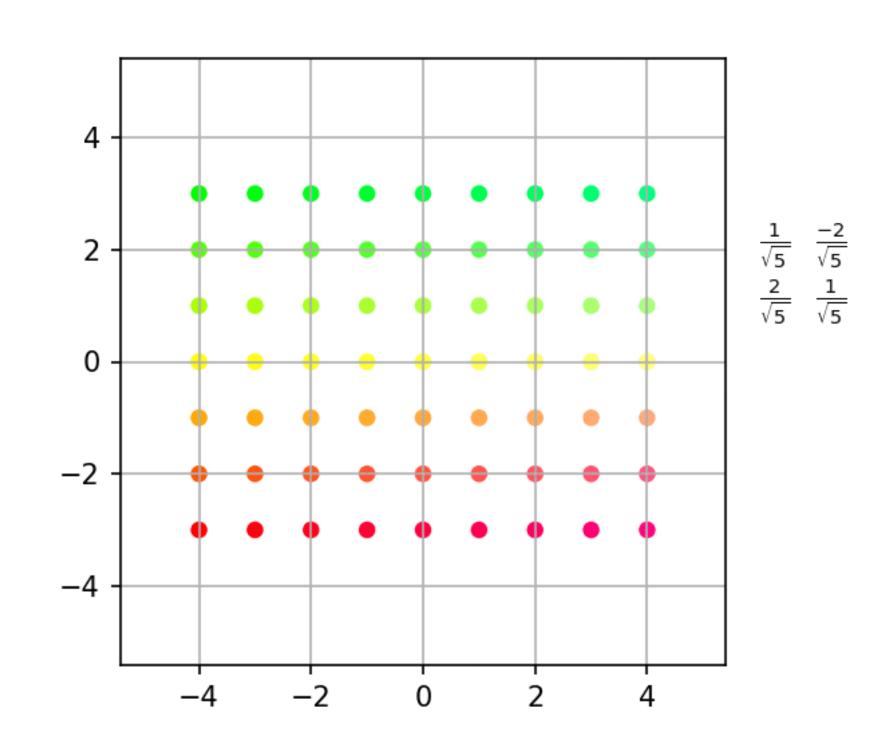
A square matrix A such that $AA^{\top} = I$ and $A^{\top}A = I$ (implies $A^{-1} = A^{\top}$)

$$A^{\mathsf{T}}A = \begin{pmatrix} \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot n} \\ \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot n} \\ \vdots & \vdots & \ddots & \vdots \\ \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot n} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

The columns of A form an orthonormal basis

Determinant of 1:
$$A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

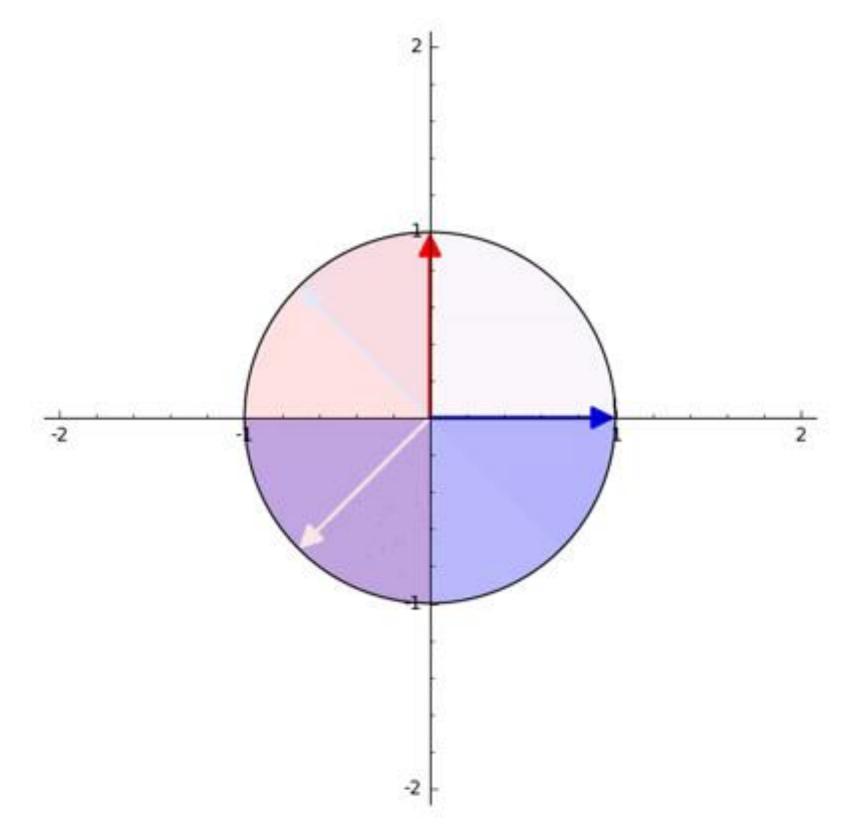


A square matrix A such that $AA^{\top} = I$ and $A^{\top}A = I$ (implies $A^{-1} = A^{\top}$)

$$A^{\mathsf{T}}A = \begin{pmatrix} \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot n} \\ \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot n} \\ \vdots & \vdots & \ddots & \vdots \\ \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot n} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Determinant of 1: $A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$



The columns of A form an orthonormal basis

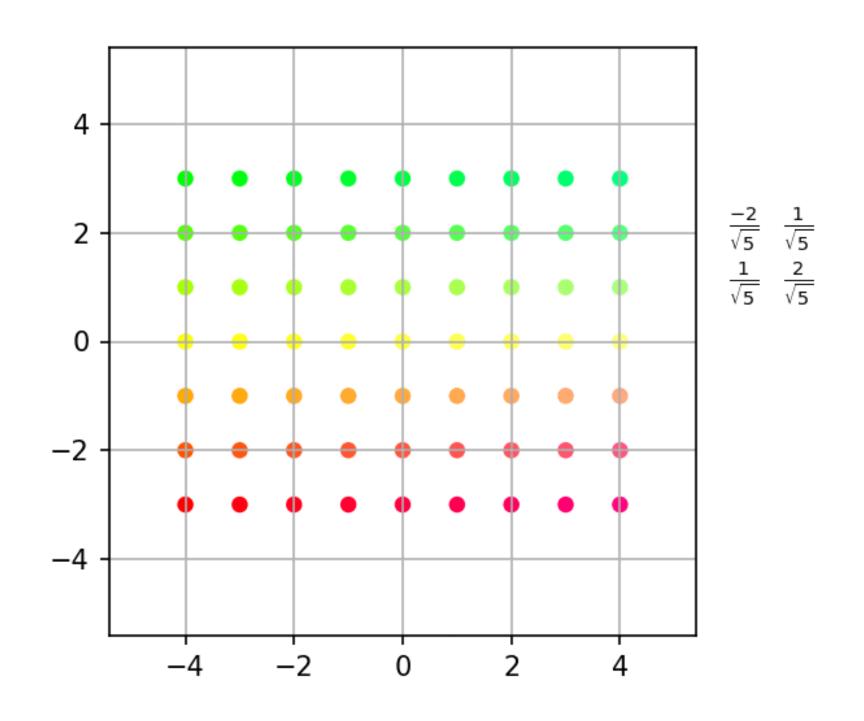
Credit: Ryan Holbrook

A square matrix A such that $AA^{\top} = I$ and $A^{\top}A = I$ (implies $A^{-1} = A^{\top}$)

$$A^{\mathsf{T}}A = \begin{pmatrix} \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot n} \\ \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot n} \\ \vdots & \vdots & \ddots & \vdots \\ \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot n} \end{pmatrix}$$

$$=egin{pmatrix} 1 & 0 & \cdots & 0 \ 0 & 1 & \cdots & 0 \ dots & dots & dots \ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Determinant of -1: $A = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$



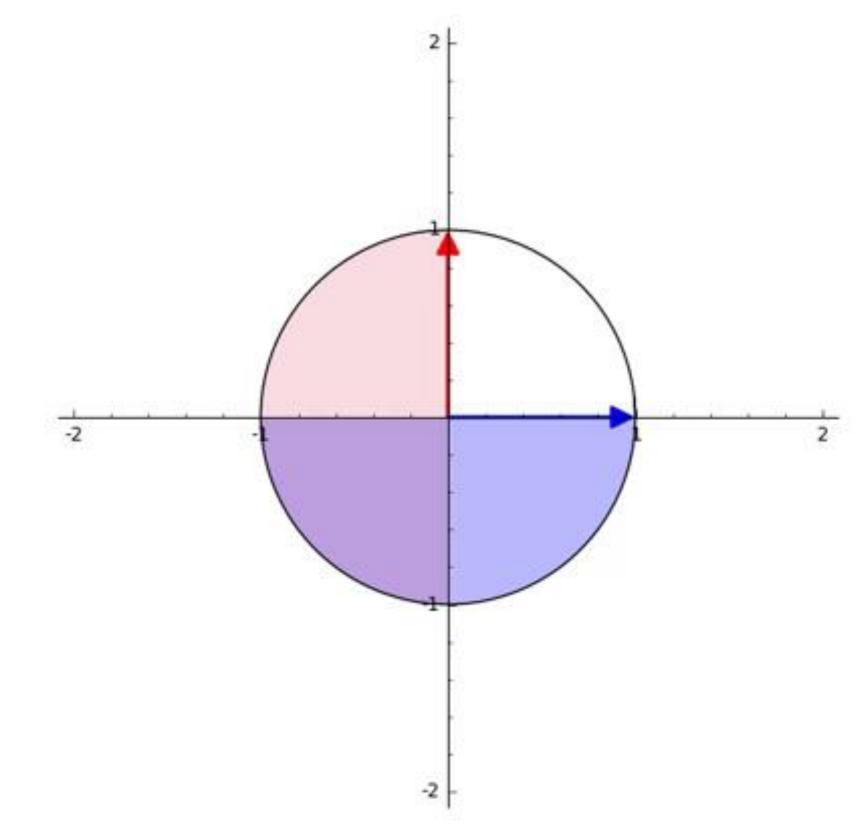
The columns of A form an orthonormal basis

A square matrix A such that $AA^{\top} = I$ and $A^{\top}A = I$ (implies $A^{-1} = A^{\top}$)

$$A^{\mathsf{T}}A = \begin{pmatrix} \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot 1}^{\mathsf{T}} \vec{a}_{\cdot n} \\ \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot 2}^{\mathsf{T}} \vec{a}_{\cdot n} \\ \vdots & \vdots & \ddots & \vdots \\ \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot 1} & \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot 2} & \cdots & \vec{a}_{\cdot n}^{\mathsf{T}} \vec{a}_{\cdot n} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Determinant of -1: $A = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}$



The columns of A form an orthonormal basis

Credit: Ryan Holbrook

Addendum (Coordinates for Non-Orthogonal Basis)

Suppose $\vec{x} = \sum_{i=1}^{N} \alpha_i \vec{v}_i$. Given \vec{x} and the basis $\{\vec{v}_i\}$, we'd like to determine the

coordinates $\{\alpha_i\}$. The sum can be written as $\vec{x} = [\vec{v}_1 \quad \cdots \quad \vec{v}_N] \begin{bmatrix} \vec{v}_1 \\ \vdots \\ \alpha_N \end{bmatrix}$

(If we let
$$v_i = \begin{bmatrix} v_{i1} \\ \vdots \\ v_{iN} \end{bmatrix}$$
. Then, we can write $\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} v_{11} & \cdots & v_{N1} \\ \vdots & \ddots & \vdots \\ v_{1N} & \cdots & v_{NN} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix}$)

Next, we'd solve the above equation for the unknowns $\{\alpha_i\}$. There are many ways. One (potentially inefficient) way is by finding the inverse:

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix} = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_N \end{bmatrix}^{-1} \vec{x}$$