Module 6.2 : Linear Algebra - Basic Definitions

- We will see some more examples where eigenvectors are important, but before that let's revisit some basic definitions from linear algebra.

Basis

A set of vectors $\in \mathbb{R}^n$ is called a basis, if they are linearly independent and every vector $\in \mathbb{R}^n$ can be expressed as a linear combination of these vectors.

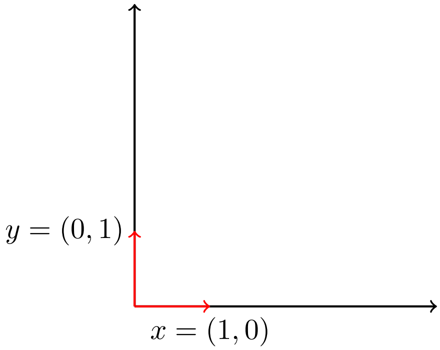
Basis

A set of vectors $\in \mathbb{R}^n$ is called a basis, if they are linearly independent and every vector $\in \mathbb{R}^n$ can be expressed as a linear combination of these vectors.

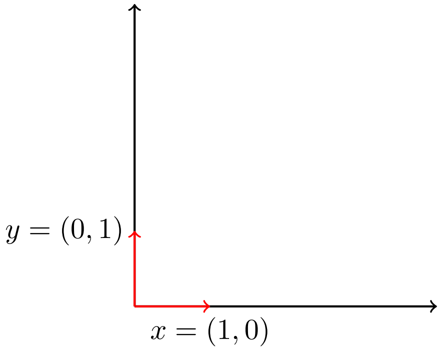
Linearly independent vectors

A set of n vectors $v_1, v_2, \ldots, v_n$ is linearly independent if no vector in the set can be expressed as a linear combination of the remaining $n-1$ vectors.
In other words, the only solution to

$$c_1 v_1 + c_2 v_2 + \ldots c_n v_n = 0 \text{ is } c_1 = c_2 = \cdots = c_n = 0 (c_i\text{'s } \textit{are scalars})$$

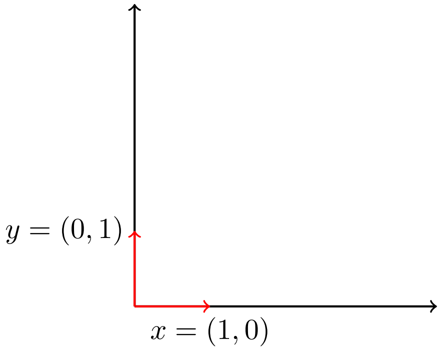


- For example consider the space $\mathbb{R}^2$



- For example consider the space $\mathbb{R}^2$
- Now consider the vectors

$$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

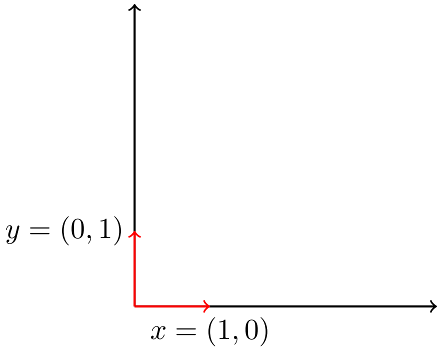


- For example consider the space $\mathbb{R}^2$
- Now consider the vectors

$$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- Any vector $\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$, can be expressed as a linear combination of these two vectors i.e

$$\begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



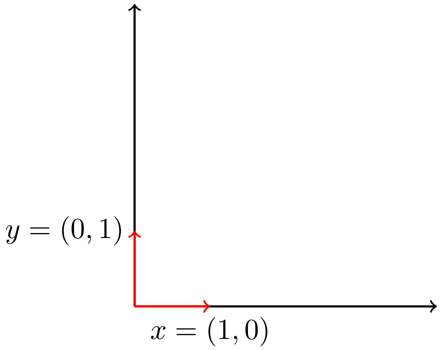
- For example consider the space $\mathbb{R}^2$
- Now consider the vectors

$$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

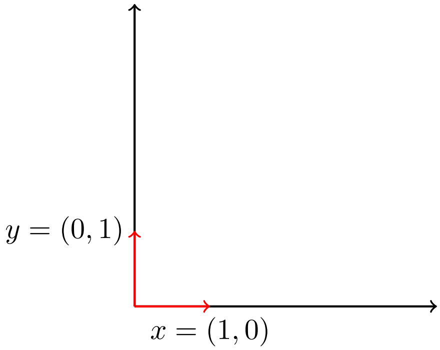
- Any vector $\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$, can be expressed as a linear combination of these two vectors i.e

$$\begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

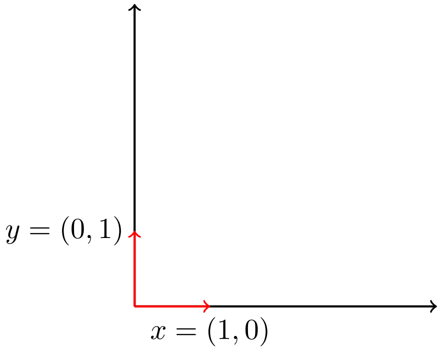
- Further, x and y are linearly independent. (the only solution to $c_1x + c_2y = 0$ is $c_1 = c_2 = 0$)



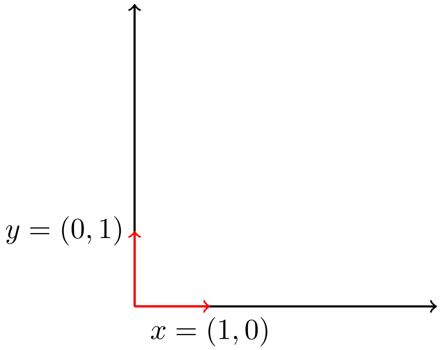
- In fact, turns out that x and y are unit vectors in the direction of the co-ordinate axes.



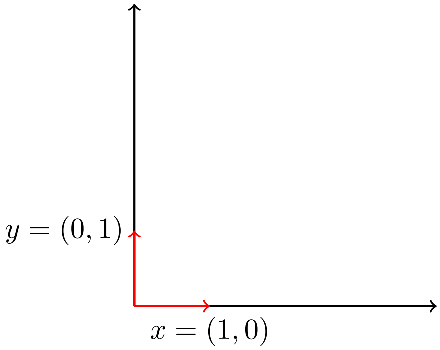
- In fact, turns out that x and y are unit vectors in the direction of the co-ordinate axes.
- And indeed we are used to representing all vectors in $\mathbb{R}^2$ as a linear combination of these two vectors.



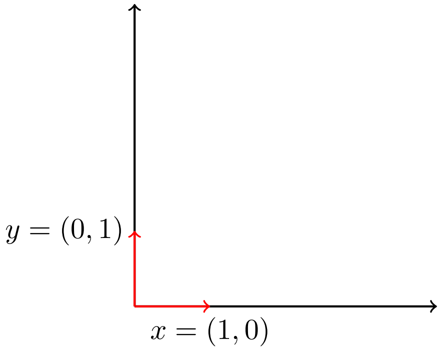
- In fact, turns out that x and y are unit vectors in the direction of the co-ordinate axes.
- And indeed we are used to representing all vectors in $\mathbb{R}^2$ as a linear combination of these two vectors.
- But there is nothing sacrosanct about the particular choice of x and y.



- In fact, turns out that x and y are unit vectors in the direction of the co-ordinate axes.
- And indeed we are used to representing all vectors in $\mathbb{R}^2$ as a linear combination of these two vectors.
- But there is nothing sacrosanct about the particular choice of x and y.
- We could have chosen any 2 linearly independent vectors in $\mathbb{R}^2$ as the basis vectors.

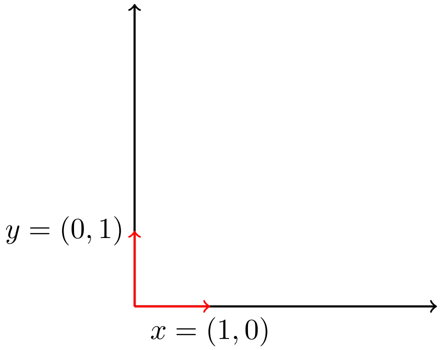


- In fact, turns out that x and y are unit vectors in the direction of the co-ordinate axes.
- And indeed we are used to representing all vectors in $\mathbb{R}^2$ as a linear combination of these two vectors.
- But there is nothing sacrosanct about the particular choice of x and y.
- We could have chosen any 2 linearly independent vectors in $\mathbb{R}^2$ as the basis vectors.
- For example, consider the linearly independent vectors, $[2, 3]^T$ and $[5, 7]^T$. See how any vector $[a, b]^T \in \mathbb{R}^2$ can be expressed as a linear combination of these two vectors.



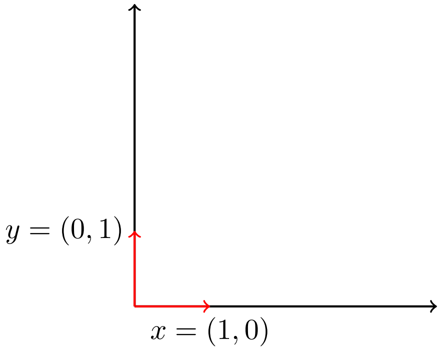
$$\begin{bmatrix} a \\ b \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

- In fact, turns out that x and y are unit vectors in the direction of the co-ordinate axes.
- And indeed we are used to representing all vectors in $\mathbb{R}^2$ as a linear combination of these two vectors.
- But there is nothing sacrosanct about the particular choice of x and y.
- We could have chosen any 2 linearly independent vectors in $\mathbb{R}^2$ as the basis vectors.
- For example, consider the linearly independent vectors, $[2, 3]^T$ and $[5, 7]^T$. See how any vector $[a, b]^T \in \mathbb{R}^2$ can be expressed as a linear combination of these two vectors.



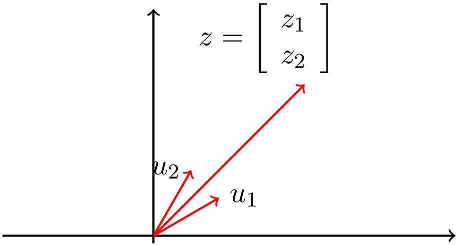
$$\begin{bmatrix} a \\ b \end{bmatrix} = x_1 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

- In fact, turns out that x and y are unit vectors in the direction of the co-ordinate axes.
- And indeed we are used to representing all vectors in $\mathbb{R}^2$ as a linear combination of these two vectors.
- But there is nothing sacrosanct about the particular choice of x and y.
- We could have chosen any 2 linearly independent vectors in $\mathbb{R}^2$ as the basis vectors.
- For example, consider the linearly independent vectors, $[2, 3]^T$ and $[5, 7]^T$. See how any vector $[a, b]^T \in \mathbb{R}^2$ can be expressed as a linear combination of these two vectors.
- We can find x_1 and x_2 by solving a system of linear equations.

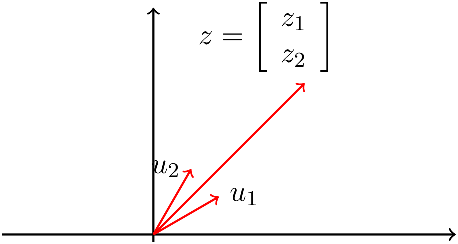


$$a = 2x_1 + 5x_2$$
$$b = 3x_1 + 7x_2$$

- In fact, turns out that x and y are unit vectors in the direction of the co-ordinate axes.
- And indeed we are used to representing all vectors in $\mathbb{R}^2$ as a linear combination of these two vectors.
- But there is nothing sacrosanct about the particular choice of x and y.
- We could have chosen any 2 linearly independent vectors in $\mathbb{R}^2$ as the basis vectors.
- For example, consider the linearly independent vectors, $[2, 3]^T$ and $[5, 7]^T$. See how any vector $[a, b]^T \in \mathbb{R}^2$ can be expressed as a linear combination of these two vectors.
- We can find x_1 and x_2 by solving a system of linear equations.

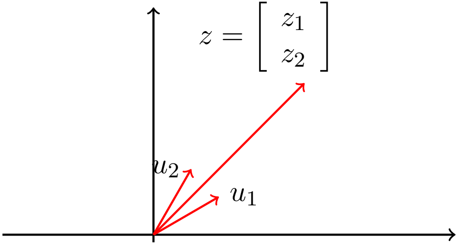


- In general, given a set of linearly independent vectors $u_1, u_2, \ldots u_n \in \mathbb{R}^n$, we can express any vector $z \in \mathbb{R}^n$ as a linear combination of these vectors.



- In general, given a set of linearly independent vectors $u_1, u_2, \ldots u_n \in \mathbb{R}^n$, we can express any vector $z \in \mathbb{R}^n$ as a linear combination of these vectors.

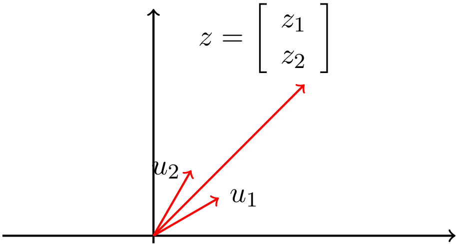
$$z = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n$$



- In general, given a set of linearly independent vectors $u_1, u_2, \dots u_n \in \mathbb{R}^n$, we can express any vector $z \in \mathbb{R}^n$ as a linear combination of these vectors.

$$z = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n$$

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \alpha_1 \begin{bmatrix} u_{11} \\ u_{12} \\ \vdots \\ u_{1n} \end{bmatrix} + \alpha_2 \begin{bmatrix} u_{21} \\ u_{22} \\ \vdots \\ u_{2n} \end{bmatrix} + \dots + \alpha_n \begin{bmatrix} u_{n1} \\ u_{n2} \\ \vdots \\ u_{nn} \end{bmatrix}$$



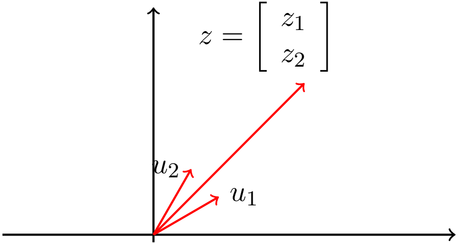
- In general, given a set of linearly independent vectors $u_1, u_2, \ldots u_n \in \mathbb{R}^n$, we can express any vector $z \in \mathbb{R}^n$ as a linear combination of these vectors.

$$z = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n$$

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \alpha_1 \begin{bmatrix} u_{11} \\ u_{12} \\ \vdots \\ u_{1n} \end{bmatrix} + \alpha_2 \begin{bmatrix} u_{21} \\ u_{22} \\ \vdots \\ u_{2n} \end{bmatrix} + \ldots + \alpha_n \begin{bmatrix} u_{n1} \\ u_{n2} \\ \vdots \\ u_{nn} \end{bmatrix}$$

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} u_{11} & u_{21} & \ldots & u_{n1} \\ u_{12} & u_{22} & \ldots & u_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ u_{1n} & u_{2n} & \ldots & u_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

(Basically rewriting in matrix form)



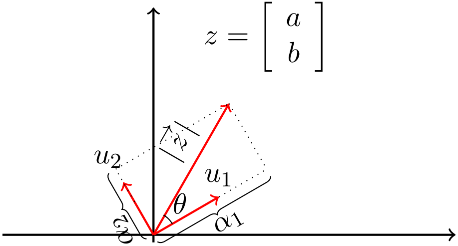
- In general, given a set of linearly independent vectors $u_1, u_2, \dots u_n \in \mathbb{R}^n$, we can express any vector $z \in \mathbb{R}^n$ as a linear combination of these vectors.

$$z = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n$$

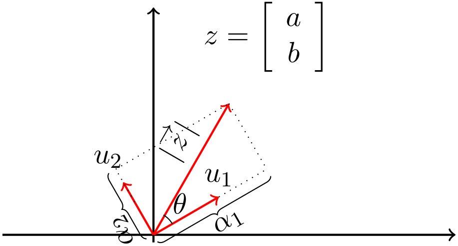
$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \alpha_1 \begin{bmatrix} u_{11} \\ u_{12} \\ \vdots \\ u_{1n} \end{bmatrix} + \alpha_2 \begin{bmatrix} u_{21} \\ u_{22} \\ \vdots \\ u_{2n} \end{bmatrix} + \ldots + \alpha_n \begin{bmatrix} u_{n1} \\ u_{n2} \\ \vdots \\ u_{nn} \end{bmatrix}$$

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} u_{11} & u_{21} & \ldots & u_{n1} \\ u_{12} & u_{22} & \ldots & u_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ u_{1n} & u_{2n} & \ldots & u_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

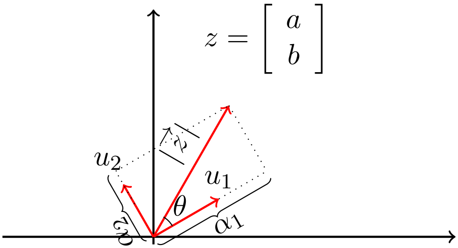
- We can now find the α_is using Gaussian Elimination (Time Complexity: $O(n^3)$)



- Now let us see if we have orthonormal basis.

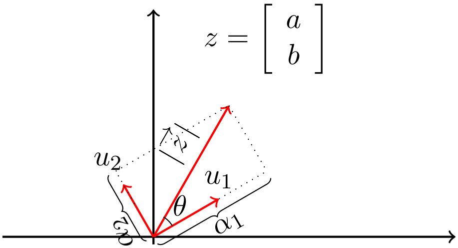


- Now let us see if we have orthonormal basis.
- $u_i^T u_j = 0 \ \forall i \neq j$ and $u_i^T u_i = \|u_i\|^2 = 1$



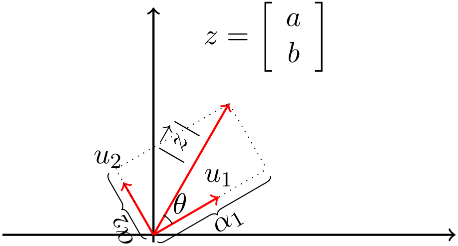
- Now let us see if we have orthonormal basis.
- $u_i^T u_j = 0 \; \forall i \neq j$ and $u_i^T u_i = \|u_i\|^2 = 1$
- Again we have:

$$z = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n$$



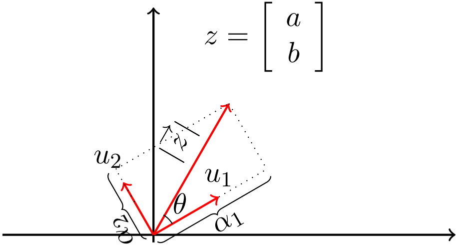
- Now let us see if we have orthonormal basis.
- $u_i^T u_j = 0 \;\forall i \neq j$ and $u_i^T u_i = \|u_i\|^2 = 1$
- Again we have:

$$z = \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n$$
$$u_1^T z = \alpha_1 u_1^T u_1 + \ldots + \alpha_n u_1^T u_n$$



- Now let us see if we have orthonormal basis.
- $u_i^T u_j = 0 \; \forall i \neq j$ and $u_i^T u_i = \|u_i\|^2 = 1$
- Again we have:

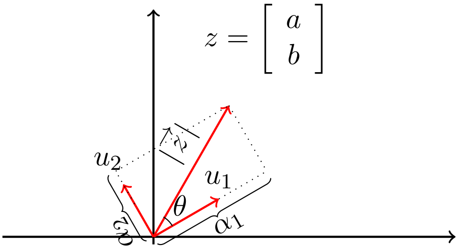
$$
\begin{aligned}
z &= \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n \\
u_1^T z &= \alpha_1 u_1^T u_1 + \ldots + \alpha_n u_1^T u_n \\
&= \alpha_1
\end{aligned}
$$



- Now let us see if we have orthonormal basis.
- $u_i^T u_j = 0 \; \forall i \neq j$ and $u_i^T u_i = \|u_i\|^2 = 1$
- Again we have:

$$\begin{aligned} z &= \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n \\ u_1^T z &= \alpha_1 u_1^T u_1 + \ldots + \alpha_n u_1^T u_n \\ &= \alpha_1 \end{aligned}$$

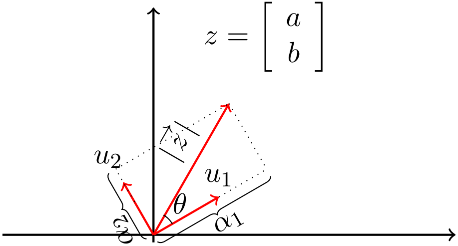
- We can directly find each α_i using a dot product between z and u_i (time complexity $O(N)$)



- Now let us see if we have orthonormal basis.
- $u_i^T u_j = 0 \; \forall i \neq j$ and $u_i^T u_i = \|u_i\|^2 = 1$
- Again we have:

$$\begin{aligned} z &= \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n \\ u_1^T z &= \alpha_1 u_1^T u_1 + \ldots + \alpha_n u_1^T u_n \\ &= \alpha_1 \end{aligned}$$

- We can directly find each α_i using a dot product between z and u_i (time complexity $O(N)$)
- The total complexity will be $O(N^2)$

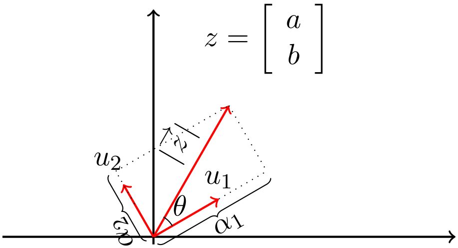


$$\alpha_1 = |\vec{z}| cos\theta = |\vec{z}| \frac{z^T u_1}{|\vec{z}||u_1|} = z^T u_1$$

- Now let us see if we have orthonormal basis.
- $u_i^T u_j = 0 \ \forall i \neq j$ and $u_i^T u_i = \|u_i\|^2 = 1$
- Again we have:

$$\begin{aligned} z &= \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n \\ u_1^T z &= \alpha_1 u_1^T u_1 + \ldots + \alpha_n u_1^T u_n \\ &= \alpha_1 \end{aligned}$$

- We can directly find each α_i using a dot product between z and u_i (time complexity $O(N)$)
- The total complexity will be $O(N^2)$



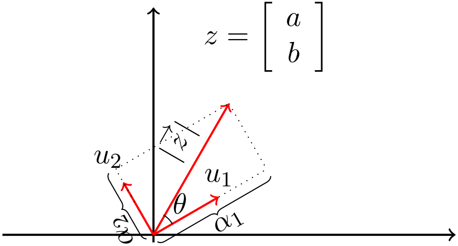
$$\alpha_1 = |\vec{z}|cos\theta = |\vec{z}|\frac{z^T u_1}{|\vec{z}||u_1|} = z^T u_1$$

Similarly, $\alpha_2 = z^T u_2$.

- Now let us see if we have orthonormal basis.
- $u_i^T u_j = 0 \; \forall i \neq j$ and $u_i^T u_i = \|u_i\|^2 = 1$
- Again we have:

$$\begin{aligned} z &= \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n \\ u_1^T z &= \alpha_1 u_1^T u_1 + \ldots + \alpha_n u_1^T u_n \\ &= \alpha_1 \end{aligned}$$

- We can directly find each α_i using a dot product between z and u_i (time complexity $O(N)$)
- The total complexity will be $O(N^2)$



$$\alpha_1 = |\vec{z}|cos\theta = |\vec{z}|\frac{z^T u_1}{|\vec{z}||u_1|} = z^T u_1$$

Similarly, $\alpha_2 = z^T u_2$.
When u_1 and u_2 are unit vectors along the co-ordinate axes

$$z = \begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- Now let us see if we have orthonormal basis.
- $u_i^T u_j = 0 \; \forall i \neq j$ and $u_i^T u_i = \|u_i\|^2 = 1$
- Again we have:

$$\begin{aligned} z &= \alpha_1 u_1 + \alpha_2 u_2 + \ldots + \alpha_n u_n \\ u_1^T z &= \alpha_1 u_1^T u_1 + \ldots + \alpha_n u_1^T u_n \\ &= \alpha_1 \end{aligned}$$

- We can directly find each α_i using a dot product between z and u_i (time complexity $O(N)$)
- The total complexity will be $O(N^2)$

Remember

An orthogonal basis is the most convenient basis that one can hope for.

- But what does any of this have to do with eigenvectors?

- But what does any of this have to do with eigenvectors?
- Turns out that the eigenvectors can form a basis.

Theorem 1

The eigenvectors of a matrix $A \in \mathbb{R}^{n \times n}$ having distinct eigenvalues are linearly independent.

Proof: See here

- But what does any of this have to do with eigenvectors?
- Turns out that the eigenvectors can form a basis.

Theorem 1

The eigenvectors of a matrix $A \in \mathbb{R}^{n \times n}$ having distinct eigenvalues are linearly independent.

Proof: See here

- But what does any of this have to do with eigenvectors?
- Turns out that the eigenvectors can form a basis.
- In fact, the eigenvectors of a square symmetric matrix are even more special.

Theorem 1

The eigenvectors of a matrix $A \in \mathbb{R}^{n \times n}$ having distinct eigenvalues are linearly independent.

Proof: See here

Theorem 2

The eigenvectors of a square symmetric matrix are orthogonal.

Proof: See here

- But what does any of this have to do with eigenvectors?
- Turns out that the eigenvectors can form a basis.
- In fact, the eigenvectors of a square symmetric matrix are even more special.

Theorem 1

The eigenvectors of a matrix $A \in \mathbb{R}^{n \times n}$ having distinct eigenvalues are linearly independent.

Proof: See here

Theorem 2

The eigenvectors of a square symmetric matrix are orthogonal.

Proof: See here

- But what does any of this have to do with eigenvectors?
- Turns out that the eigenvectors can form a basis.
- In fact, the eigenvectors of a square symmetric matrix are even more special.
- Thus they form a very convenient basis.

Theorem 1

The eigenvectors of a matrix $A \in \mathbb{R}^{n \times n}$ having distinct eigenvalues are linearly independent.

Proof: See here

Theorem 2

The eigenvectors of a square symmetric matrix are orthogonal.

Proof: See here

- But what does any of this have to do with eigenvectors?
- Turns out that the eigenvectors can form a basis.
- In fact, the eigenvectors of a square symmetric matrix are even more special.
- Thus they form a very convenient basis.
- Why would we want to use the eigenvectors as a basis instead of the more natural co-ordinate axes?

Theorem 1

The eigenvectors of a matrix $A \in \mathbb{R}^{n \times n}$ having distinct eigenvalues are linearly independent.

Proof: See here

Theorem 2

The eigenvectors of a square symmetric matrix are orthogonal.

Proof: See here

- But what does any of this have to do with eigenvectors?
- Turns out that the eigenvectors can form a basis.
- In fact, the eigenvectors of a square symmetric matrix are even more special.
- Thus they form a very convenient basis.
- Why would we want to use the eigenvectors as a basis instead of the more natural co-ordinate axes?
- We will answer this question soon.