Linear Algebra: An Introduction

Subhrajyoty Roy

May 14, 2020

Motivating Requirements of Linear Algebra

A Consistency check for Linear Algebra

1. Let a, b be two vectors. (Just like you learn in physics, for now).
2. $\langle a, b \rangle = a \cdot b = |a||b| \cos \theta$, where θ is the angle between the vectors a and b.

A Consistency check for Linear Algebra

1. Let a, b be two vectors. (Just like you learn in physics, for now).
2. $\langle a, b\rangle = a \cdot b = |a||b| \cos\theta$, where θ is the angle between the vectors a and b.
3. Also, $a \cdot b = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$.

A Consistency check for Linear Algebra

1. Let a, b be two vectors. (Just like you learn in physics, for now).
2. $\langle a, b \rangle = a \cdot b = |a||b| \cos\theta$, where θ is the angle between the vectors a and b.
3. Also, $a \cdot b = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$.
4. Also, by Taylor's series expansion,

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} \ldots$$

A Consistency check for Linear Algebra

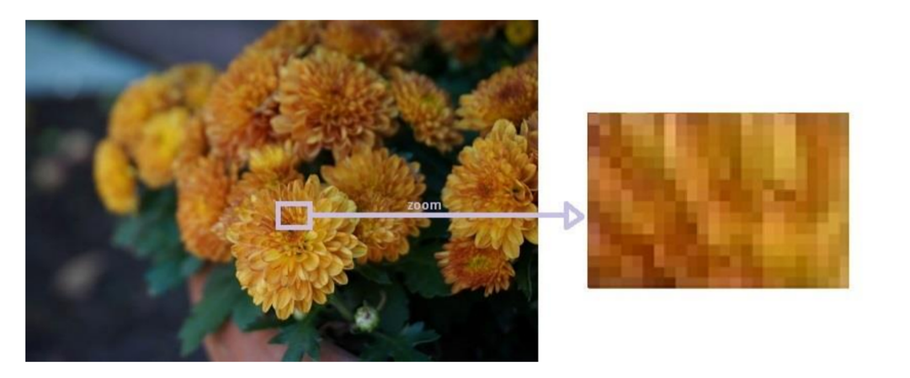
1. Let a, b be two vectors. (Just like you learn in physics, for now).
2. $\langle a, b\rangle = a \cdot b = |a||b|\cos\theta$, where θ is the angle between the vectors a and b.
3. Also, $a \cdot b = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$.
4. Also, by Taylor's series expansion,

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} \dots$$

5. So, we have;

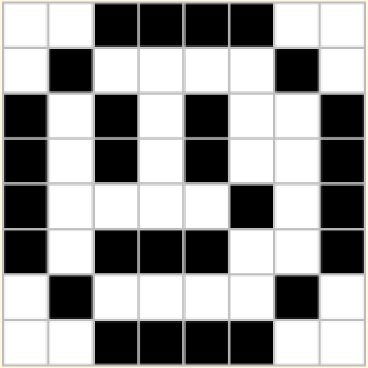
$$\left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} \dots\right) = \frac{(a_1 b_1 + a_2 b_2 + \cdots + a_n b_n)}{\sqrt{a_1^2 + \cdots + a_n^2}\sqrt{b_1^2 + \cdots + b_n^2}}$$

An Application of Linear Algebra



Reference: Digital Image Matrix Operations by Williams Orenda

An Application of Linear Algebra



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

An Application of Linear Algebra

1. A Grayscale image of 512×480 pixels is represented by a matrix of order 512×480.
2. A color image of 512×480 pixels is represented by three matrices of order 512×480. We call these matrices as channels. There are Red, Green, Blue channels.
3. Now, we do like put the third matrix at first position, and the first matrix at third position. Hence, the high values of red matrix, now denotes the high value of blue matrix.

An Application of Linear Algebra

Figure 4: red-blue swap

Reference: Digital Image Matrix Operations by Williams Orenda

Building Blocks of Linear Algebra

Five W's about Numbers

1. What is a number? Why is it the case that "1, 2, 3" are numbers, but "I" am not?

Five W's about Numbers

1. What is a number? Why is it the case that "1, 2, 3" are numbers, but "I" am not?
2. Where does number comes from?

Five W's about Numbers

1. What is a number? Why is it the case that "1, 2, 3" are numbers, but "I" am not?
2. Where does number comes from?
3. Why are numbers required at all?

Five W's about Numbers

1. What is a number? Why is it the case that "1, 2, 3" are numbers, but "I" am not?
2. Where does number comes from?
3. Why are numbers required at all?
4. How numbers solve these problems?

Five W's about Numbers

1. What is a number? Why is it the case that "1, 2, 3" are numbers, but "I" am not?
2. Where does number comes from?
3. Why are numbers required at all?
4. How numbers solve these problems?
5. When we should think about numbers?

Numbers vs Tuple of Numbers

Field $\mathcal{F}$

1. Two operations $+, \cdot$.
2. $a + b = b + a$ and $ab = ba$.
3. $a(b + c) = ab + ac$.
4. $a + (b + c) = (a + b) + c$.
5. $\exists 0$ s.t. $a + 0 = a$.
6. $\forall a \exists b$ s.t. $a + b = 0$.
7. $\exists 1$ s.t. $a \cdot 1 = a$.
8. $\forall a \neq 0 \exists b'$ s.t. $a \cdot b' = 1$.
9. $0 \neq 1$.

Numbers vs Tuple of Numbers

Field $\mathcal{F}$

1. Two operations $+, \cdot$.
2. $a + b = b + a$ and $ab = ba$.
3. $a(b + c) = ab + ac$.
4. $a + (b + c) = (a + b) + c$.
5. $\exists 0$ s.t. $a + 0 = a$.
6. $\forall a \exists b$ s.t. $a + b = 0$.
7. $\exists 1$ s.t. $a \cdot 1 = a$.
8. $\forall a \neq 0 \exists b'$ s.t. $a \cdot b' = 1$.
9. $0 \neq 1$.

Vector Space $\mathcal{V}$ over field $\mathcal{F}$

1. Two operations $\oplus, \odot$.
2. $u \oplus v = v \oplus u$.
3. $u \oplus (v \oplus w) = (u \oplus v) \oplus w$.
4. $\exists 0_{\mathcal{V}}$, s.t. $v \oplus 0_{\mathcal{V}} = v$.
5. $\forall v \exists u$ s.t. $v \oplus u = 0_{\mathcal{V}}$.
6. $\forall \alpha \in \mathcal{F}$,
   $\alpha \odot (u \oplus v) = (\alpha \odot u) \oplus (\alpha \odot v)$.
7. $\forall \alpha, \beta \in \mathcal{F}$,
   $(\alpha + \beta) \odot v = (\alpha \odot v) \oplus (\beta \odot v)$.
8. $(\alpha\beta) \odot v = \alpha \odot (\beta \odot v)$.
9. $1_{\mathcal{F}} \odot v = v$.

Spanning set in 2D Vector Space

1. There is a special car, which goes only up-down.

$$\downarrow\uparrow$$

Can you reach anywhere on the 2d plane with it?

Spanning set in 2D Vector Space

1. There is a special car, which goes only up-down.

$$\downarrow\!\uparrow$$

   Can you reach anywhere on the 2d plane with it?

2. The car goes up-down, and also left-right.

$$\downarrow\!\uparrow\rightleftarrows$$

   Can you reach anywhere on the 2d plane with it?

Spanning set in 2D Vector Space

1. There is a special car, which goes only up-down.

   $$\downarrow\uparrow$$

   Can you reach anywhere on the 2d plane with it?

2. The car goes up-down, and also left-right.

   $$\downarrow\uparrow\rightleftarrows$$

   Can you reach anywhere on the 2d plane with it?

3. The car goes up-down, left-right, and also cornerwise.

   $$\downarrow\uparrow\rightleftarrows\nearrow\swarrow\ \nwarrow\searrow$$

   Can you reach anywhere on the 2d plane with it?

Spanning Set

Definition

A set of vectors $\{v_1, v_2, \ldots v_k\}$ is said to span the whole vector space $\mathcal{V}$, if for any vector $v \in \mathcal{V}$, there exists some scalers $\alpha_1, \alpha_2, \ldots \alpha_k \in \mathcal{F}$, such that,

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$$

This is called a linear combination.
In the previous example, the first set of vectors was not spanning, but the second and third was.

Linear Independence

Consider a bird returning to its home. It visualizes the world as a 3d space.[1] For now, let x is the axis straightways, y is the axis sideways, and z is the axis from top of the sky to bottom of the ground.

1. Suppose the bird knows how to fly straight and sideways. Can it fly downwards or upwards?

[1] We shall be more clear about what we mean by 3d or 3 dimension

Linear Independence

Consider a bird returning to its home. It visualizes the world as a 3d space.[1] For now, let x is the axis straightways, y is the axis sideways, and z is the axis from top of the sky to bottom of the ground.

1. Suppose the bird knows how to fly straight and sideways. Can it fly downwards or upwards?
2. Suppose the bird knows how to fly straight and sideways. Can it fly parallel $x = y$ plane?

[1] We shall be more clear about what we mean by 3d or 3 dimension

Linearly Independent Set

Definition

A set of vectors $\{v_1, v_2, \dots v_k\}$ is said to be linearly independent, if for any v_i, cannot be written as a linear combination of the other vectors in the set, i.e. one cannot find some scalers $\alpha_1, \dots \alpha_{i-1}, \alpha_{i+1}, \dots \alpha_k \in \mathcal{F}$ such that,

$$v_i = \alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k$$

Linearly Independent Set

Definition

A set of vectors $\{v_1, v_2, \dots v_k\}$ is said to be linearly independent, if for any v_i, cannot be written as a linear combination of the other vectors in the set, i.e. one cannot find some scalers $\alpha_1, \dots \alpha_{i-1}, \alpha_{i+1}, \dots \alpha_k \in \mathcal{F}$ such that,

$$v_i = \alpha_1 v_1 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + \alpha_k v_k$$

OR A set of vectors $\{v_1, v_2, \dots v_k\}$ is said to be linearly independent if there does not exist some scalers $\alpha_1, \alpha_2, \dots \alpha_k \in \mathcal{F}$, such that,

$$0 = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$$

where atleast one $\alpha_i \neq 0$.

Two Basic Results

Result (Extension of Spanning Set)

If S is a spanning set, then any superset of S is also a spanning set.

Two Basic Results

Result (Extension of Spanning Set)

If S is a spanning set, then any superset of S is also a spanning set.

Result (Deduction of Independent Set)

If S is a linearly independent set, then any subset of S is also a linearly independent set.

Basis of a Vector Space

Definition

A basis of $\mathcal{V}$ is a spanning set for $\mathcal{V}$ which is also linearly independent.

Exercise

1. *Show that, a basis of $\mathcal{V}$ is the smallest possible spanning set of $\mathcal{V}$.*
2. *Show that, a basis is the largest possible linearly independent set.*

The dimension of vector space

Theorem

If B_1 and B_2 are two bases of a vector space $\mathcal{V}$, then $|B_1| = |B_2|$, i.e. their sizes are same.

The dimension of vector space

Theorem

If B_1 and B_2 are two bases of a vector space $\mathcal{V}$, then $|B_1| = |B_2|$, i.e. their sizes are same.

To prove this, let, $B_1 = \{v_1, v_2, \ldots v_m\}$ and $B_2 = \{u_1, u_2, \ldots u_n\}$.

The dimension of vector space

Theorem

If B_1 and B_2 are two bases of a vector space $\mathcal{V}$, then $|B_1| = |B_2|$, i.e. their sizes are same.

To prove this, let, $B_1 = \{v_1, v_2, \ldots v_m\}$ and $B_2 = \{u_1, u_2, \ldots u_n\}$. However, We need another lemma.

Lemma (Replacement Theorem)

There exists v_i such that, $B_1^{(1)} = \{u_1, v_1, v_2, \ldots v_{i-1}, v_{i+1}, \ldots v_m\}$, which is obtained by including u_1 and removing v_i, is also a basis.

Proof of Basis Theorem

Proof.

Assume $m \neq n$, we shall use proof by contradiction. Without loss of generality, $m < n$.
Consider, replacing u_1 to get, $B_1^{(1)}$, which is a basis. Next, replace u_2 to $B_1^{(1)}$ to get the new basis, $B_1^{(2)}$.
Continue replacing all elements from B_2, until $B_1^{(k)}$ is completely filled with u_i's only, (i.e. all v_i's are removed).
But, this final basis is maximal linear independent set, but is a strict subset of B_2. B_2 cannot be basis. □

Example 1

Vector space $\mathbb{R}^n$, which is the space of n-dimensional vector.

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

is the Euclidean basis.
To express an element $x \in \mathbb{R}^n$, we use;

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \end{pmatrix}$$

Identify x as the tuple of the coefficients $(x_1, x_2, \dots x_n)$.

Example 2

Vector space $\mathbb{R}^2$, which is the space of 2 dimensional vector with real elements.

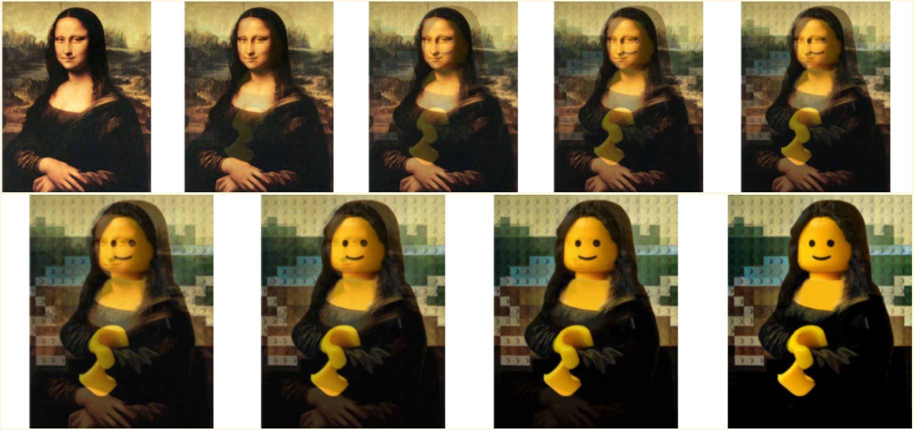
$$B = \left\{ \left(1, 1\right), \left(1, 2\right) \right\}$$

is a basis.
To express the usual vector $x = (4, 5)$, with respect to this basis,

$$x = \begin{pmatrix} 4 \\ 5 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

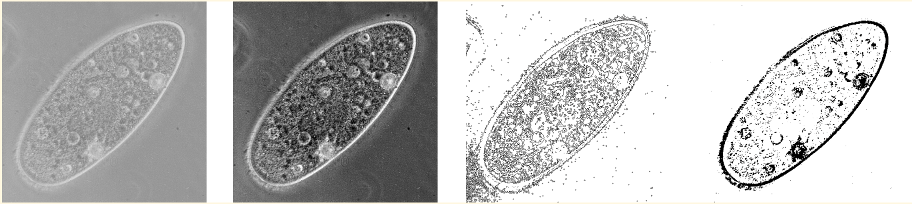
So identify x as $(3, 1)$, the tuple of coefficients.

Application of taking Linear Combinations



Matrix as a Linear Transformation

Application of identifying matrix as function



What is Linear Transformation

Definition

A linear transformation is a function f from a vector space $\mathcal{U}$ to a vector space $\mathcal{V}$, such that;

What is Linear Transformation

Definition

A linear transformation is a function f from a vector space $\mathcal{U}$ to a vector space $\mathcal{V}$, such that;

1. $\mathcal{U}$ and $\mathcal{V}$ are both vector spaces over same field $\mathcal{F}$.

What is Linear Transformation

Definition

A linear transformation is a function f from a vector space $\mathcal{U}$ to a vector space $\mathcal{V}$, such that;

1. $\mathcal{U}$ and $\mathcal{V}$ are both vector spaces over same field $\mathcal{F}$.
2. $f(\alpha v_1 \oplus \beta v_2) = \alpha f(v_1) \oplus \beta f(v_2)$, for any $\alpha, \beta \in \mathcal{F}$ and $v_1, v_2 \in \mathcal{U}$.

Examples of Linear Transformation

Example

Let $f : \mathbb{R}^n \to \mathbb{R}^2$ be a function such that,

$$f(v) = \begin{cases} (1,0) & \text{if there are odd number of positive entries in v} \\ (0,1) & \text{otherwise} \end{cases}$$

Is it a linear transformation?

Examples of Linear Transformation

Example

Let $f : \mathbb{R}^n \to \mathbb{R}^2$ be a function such that,

$$f(v) = \begin{cases} (1,0) & \text{if there are odd number of positive entries in v} \\ (0,1) & \text{otherwise} \end{cases}$$

Is it a linear transformation?

Example

Let $f : \mathbb{R}^n \to \mathbb{R}^2$ be a function such that,

$$f(v) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

where x_1 is sum of odd position elements, x_2 is sum of even position elements. Is it a linear transformation?

What's so good about it?

Let, $f : \mathcal{U} \to \mathcal{V}$ be a linear transformation. Let, $A = \{u_1, u_2, \dots u_m\}$ be the basis of $\mathcal{U}$.
Then,

$$u \in \mathcal{U} \implies u = \sum_i \alpha_i u_i$$

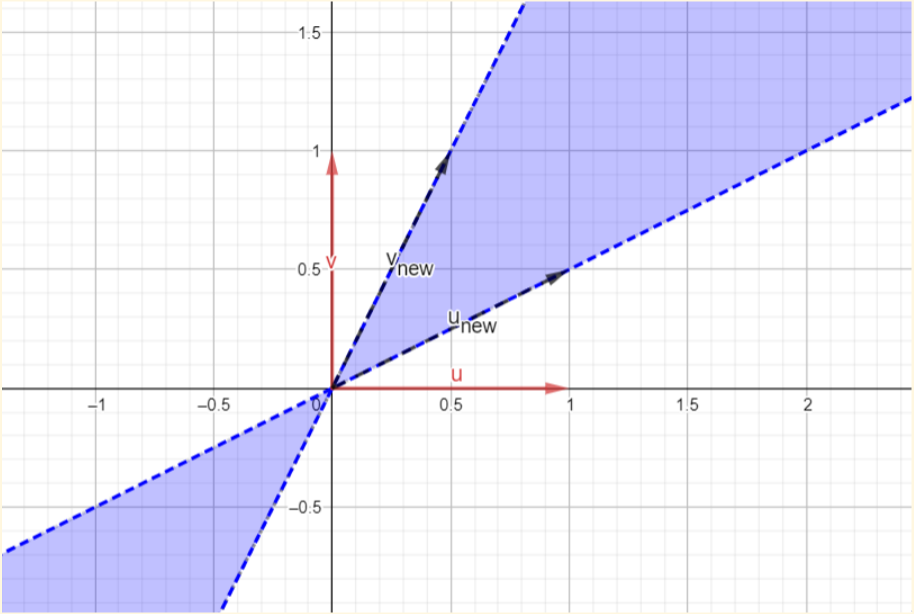
So,

$$f(u) = f\left(\sum_i \alpha_i u_i\right) = \sum_i \alpha_i f(u_i)$$

Result

To specify a linear transformation, it is just enough to know its action on the basis.

Example



Linear Transformation to Matrix

Now, let $B = \{v_1, v_2, \ldots v_n\}$ be a basis of vector space $\mathcal{V}$. Then, there are scalars, r_{ij}, $i = 1, 2, \ldots m$ and $j = 1, 2, \ldots n$ such that,

Linear Transformation to Matrix

Now, let $B = \{v_1, v_2, \dots v_n\}$ be a basis of vector space $\mathcal{V}$. Then, there are scalars, r_{ij}, $i = 1, 2, \dots m$ and $j = 1, 2, \dots n$ such that,

$$
\begin{aligned}
f(u_1) &= r_{11} v_1 + r_{12} v_2 + \cdots + r_{1n} v_n \\
f(u_2) &= r_{21} v_1 + r_{22} v_2 + \cdots + r_{2n} v_n \\
\dots\dots \\
f(u_m) &= r_{m1} v_1 + r_{m2} v_2 + \cdots + r_{mn} v_n
\end{aligned}
$$

We can now collect all these numbers r_{ij} together, so that we get an array of scalars, with m rows and n columns. This is **MATRIX** over the scalar field $\mathcal{F}$, corresponding to the transformation $f : U \to V$ with respect to the basis A and B.

Generalizing Matrix Vector Product

1. Take $u \in \mathcal{U}$, so we have $u = \sum_i \alpha_i u_i$.

Generalizing Matrix Vector Product

1. Take $u \in \mathcal{U}$, so we have $u = \sum_i \alpha_i u_i$.
2. Now, $f(u) = \sum_i \alpha_i f(u_i)$.

Generalizing Matrix Vector Product

1. Take $u \in \mathcal{U}$, so we have $u = \sum_i \alpha_i u_i$.
2. Now, $f(u) = \sum_i \alpha_i f(u_i)$.
3. $f(u) = \sum_i \alpha_i \sum_j r_{ij} v_j = \sum_j \sum_i r_{ij} \alpha_i$

Generalizing Matrix Vector Product

1. Take $u \in \mathcal{U}$, so we have $u = \sum_i \alpha_i u_i$.
2. Now, $f(u) = \sum_i \alpha_i f(u_i)$.
3. $f(u) = \sum_i \alpha_i \sum_j r_{ij} v_j = \sum_j \sum_i r_{ij} \alpha_i$
4. But we identify a vector by its coefficients with respect to the basis.

Generalizing Matrix Vector Product

1. Take $u \in \mathcal{U}$, so we have $u = \sum_i \alpha_i u_i$.
2. Now, $f(u) = \sum_i \alpha_i f(u_i)$.
3. $f(u) = \sum_i \alpha_i \sum_j r_{ij} v_j = \sum_j \sum_i r_{ij} \alpha_i$
4. But we identify a vector by its coefficients with respect to the basis.
5. So, we have the following transformation;

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_m \end{pmatrix} \to \begin{pmatrix} \sum_i r_{i1}\alpha_i \\ \sum_i r_{i2}\alpha_i \\ \dots \\ \sum_i r_{in}\alpha_i \end{pmatrix} = \begin{bmatrix} r_{11} & r_{21} & \dots & r_{m1} \\ r_{12} & r_{22} & \dots & r_{m2} \\ \vdots & \ddots & \vdots & \vdots \\ r_{1n} & r_{2n} & \dots & r_{mn} \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_m \end{pmatrix}$$

Generalizing Matrix Vector Product

1. Take $u \in \mathcal{U}$, so we have $u = \sum_i \alpha_i u_i$.
2. Now, $f(u) = \sum_i \alpha_i f(u_i)$.
3. $f(u) = \sum_i \alpha_i \sum_j r_{ij} v_j = \sum_j \sum_i r_{ij} \alpha_i$
4. But we identify a vector by its coefficients with respect to the basis.
5. So, we have the following transformation;

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_m \end{pmatrix} \rightarrow \begin{pmatrix} \sum_i r_{i1}\alpha_i \\ \sum_i r_{i2}\alpha_i \\ \dots \\ \sum_i r_{in}\alpha_i \end{pmatrix} = \begin{bmatrix} r_{11} & r_{21} & \dots & r_{m1} \\ r_{12} & r_{22} & \dots & r_{m2} \\ \vdots & \ddots & \vdots & \vdots \\ r_{1n} & r_{2n} & \dots & r_{mn} \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_m \end{pmatrix}$$

6. $f(u_{m\times 1})_{n\times 1} = (R_f)_{n\times m} u_{m\times 1}$. Note the transpose of matrix.

Example 1

Example

$f : \mathbb{R}^4 \to \mathbb{R}^2$ such that,

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}\right) = \begin{pmatrix} x_1 + x_3 \\ x_2 + x_4 \end{pmatrix}$$

is given by the matrix;

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}_{2\times 4}$$

Example 2

Example

$f : \mathbb{R}^2 \to \mathbb{R}^2$ such that,

$$f(z) = ze^{i\theta}$$

where $z \in \mathbb{C}$. The transformation matrix works as follows;

$$\begin{pmatrix} a \\ b \end{pmatrix} \to \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

This type of rotational transformation is denoted by a matrix of special property, these are called **Orthogonal** matrices.

Some more results

Theorem

1. *If* f, g *are linear transformation with matrices* A *and* B*, show that,* $(f \pm g)$ *has matrix* $A \pm B$ *with respect to same bases.*

Some more results

Theorem

1. *If f, g are linear transformation with matrices A and B, show that, $(f \pm g)$ has matrix $A \pm B$ with respect to same bases.*
2. *$f \circ g$ has matrix AB, where it denotes the matrix multiplication.*

Some more results

Theorem

1. *If* f, g *are linear transformation with matrices* A *and* B*, show that,* $(f \pm g)$ *has matrix* $A \pm B$ *with respect to same bases.*
2. $f \circ g$ *has matrix* AB*, where it denotes the matrix multiplication.*
3. f^{-1}*, if exists, has the matrix* A^{-1}*, where* $AA^{-1} = I$*, the identity matrix.*

Some more results

Theorem

1. *If f, g are linear transformation with matrices A and B, show that, $(f \pm g)$ has matrix $A \pm B$ with respect to same bases.*
2. *$f \circ g$ has matrix AB, where it denotes the matrix multiplication.*
3. *f^{-1}, if exists, has the matrix A^{-1}, where $AA^{-1} = I$, the identity matrix.*

Exercise

1. *How does the transformation $f : \mathbb{R}^n \to \mathbb{R}^m$ given by $f(x) = x$ looks like as matrix? Is it identity matrix?*
2. *What is the matrix corresponding to the transformation fg, i.e. $(fg)(u) = f(u)g(u)$.*
3. *What is the matrix corresponding to f/g? Is it AB^{-1}? or $B^{-1}A$?*
4. *What is the linear transformation corresponding to the matrix A^T? What is corresponding to $A \otimes B$, where $\otimes$ is the elementwise product.*

Kernel, Null Space and Range

Definition

The kernel or null space of a transformation f is the set $Ker(f) = \{x : f(x) = 0\}$. That means, the null space of a matrix A is $\mathcal{N}(A) = \{x : Ax = 0\}$.

Kernel, Null Space and Range

Definition

The kernel or null space of a transformation f is the set $Ker(f) = \{x : f(x) = 0\}$. That means, the null space of a matrix A is $\mathcal{N}(A) = \{x : Ax = 0\}$.

Definition

The range of a transformation $f : \mathcal{U} \to \mathcal{V}$ is the set $R(f) = \{f(x) : x \in \mathcal{U}\}$.

Kernel, Null Space and Range

Definition

The kernel or null space of a transformation f is the set $Ker(f) = \{x : f(x) = 0\}$. That means, the null space of a matrix A is $\mathcal{N}(A) = \{x : Ax = 0\}$.

Definition

The range of a transformation $f : \mathcal{U} \to \mathcal{V}$ is the set $R(f) = \{f(x) : x \in \mathcal{U}\}$.

Exercise

1. *If $f(x_0) = y_0$, then all solutions to the equation $f(x) = y_0$ is the set $x_0 + Ker(f)$.*
2. *Show that, $Ker(f)$ is a vector space.*

An Interesting Example

Let there are n cities, $C_1, C_2, \ldots, C_n$. Between the cities, we have one-way roads. Consider a $n \times n$ matrix R such that,

$$R_{ij} = \begin{cases} 1 & \text{if there is road from city i to city j} \\ 0 & \text{otherwise} \end{cases}$$

Then, consider square of this matrix.

$$(R^2)_{ij} = \sum_k R_{ik} R_{kj}$$

A term in the sum is 1 iff there is a road from city i to city j, through city k. So, $(R^2)_{ij}$ is the number of ways to reach city j from city i visting a city in between.
In other words, $(R + R^2 + \cdots + R^s)_{ij}$ is the number of ways to reach city i from city j in atmost s steps.

Building Row Rank and Column Rank

Consider a matrix A of order $m \times n$.

Definition

Let $\mathcal{R}(A)$ denote the span of the row vectors in A. Then, row rank of a matrix is the dimension of that spanning vector space.

Building Row Rank and Column Rank

Consider a matrix A of order $m \times n$.

Definition

Let $\mathcal{R}(A)$ denote the span of the row vectors in A. Then, row rank of a matrix is the dimension of that spanning vector space.

Definition

Let $\mathcal{C}(A)$ denote the span of the column vectors in A. Then, column rank of a matrix is the dimension of that spanning vector space.

Some Results

Theorem

$$\textit{Row rank} \leq \min\{m, n\}$$

Some Results

Theorem

$$\textit{Row rank} \leq \min\{m, n\}$$

Theorem

$$\textit{Column rank} \leq \min\{m, n\}$$

Some Results

Theorem

$$\textit{Row rank} \leq \min\{m, n\}$$

Theorem

$$\textit{Column rank} \leq \min\{m, n\}$$

Exercise

Show that, column space of a matrix $A = \mathcal{C}(A)$ *is same as the range space* $R(f)$.

Rank Theorems

Theorem

$$\textit{Row rank} = \textit{Column Rank}$$

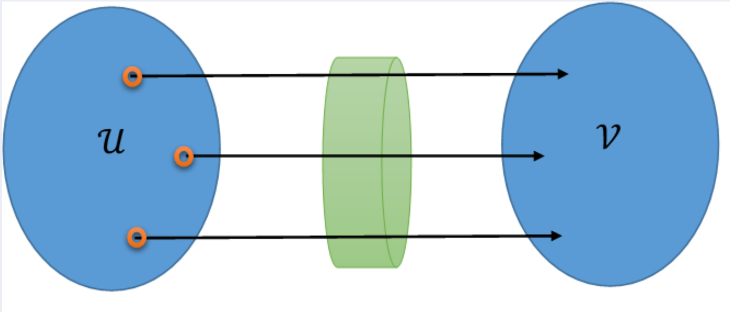
This common value is called Rank of a matrix.

Theorem

Row rank = Column Rank

This common value is called Rank of a matrix.

Proof.



□

Rank Nullity Theorem

Theorem (Rank Nullity Theorem)

If $f : \mathcal{U} \to \mathcal{V}$ is a linear transformation with corresponding matrix A, then;

$$\rho(A) + d(\mathcal{N}(A)) = d(\mathcal{V})$$

Rank Nullity Theorem

Theorem (Rank Nullity Theorem)

If $f : \mathcal{U} \to \mathcal{V}$ is a linear transformation with corresponding matrix A, then;

$$\rho(A) + d(\mathcal{N}(A)) = d(\mathcal{V})$$

Proof.

Let, $B = \{v_1, v_2, \dots v_n\}$ be a basis of $\mathcal{V}$. Some of these form the basis of $R(f)$. The rests are like useless fellows.
Enough to show, the number of such useless fellows is same as the dimension of $d(Ker(f))$ □

Invertibility and Rank of a Matrix

Theorem

A matrix is invertible if and only if it has full rank, i.e. A $n \times n$ matrix is invertible if and only if $\rho(A) = n$.

Invertibility and Rank of a Matrix

Theorem

A matrix is invertible if and only if it has full rank, i.e. A $n \times n$ matrix is invertible if and only if $\rho(A) = n$.

Proof.

Suppose, $f : \mathcal{F}^n \to \mathcal{F}^n$ is the linear transformation corresponding to A. Let, $\rho(A) < n$. That means, $d(\mathcal{R}(f)) < n$. So, there is a basis $B = \{u_1, u_2, \ldots u_{n-1}\}$ of $\mathcal{R}(f)$.

Invertibility and Rank of a Matrix

Theorem

A matrix is invertible if and only if it has full rank, i.e. A $n \times n$ matrix is invertible if and only if $\rho(A) = n$.

Proof.

Suppose, $f : \mathcal{F}^n \to \mathcal{F}^n$ is the linear transformation corresponding to A.
Let, $\rho(A) < n$. That means, $d(\mathcal{R}(f)) < n$. So, there is a basis $B = \{u_1, u_2, \ldots u_{n-1}\}$ of $\mathcal{R}(f)$.
There must be a vector $v \in \mathcal{F}^n$ such that, $B \cup \{v\}$ is linearly independent, hence is basis of $\mathcal{F}^n$.

Invertibility and Rank of a Matrix

Theorem

A matrix is invertible if and only if it has full rank, i.e. A $n \times n$ matrix is invertible if and only if $\rho(A) = n$.

Proof.

Suppose, $f : \mathcal{F}^n \to \mathcal{F}^n$ is the linear transformation corresponding to A.
Let, $\rho(A) < n$. That means, $d(\mathcal{R}(f)) < n$. So, there is a basis $B = \{u_1, u_2, \ldots u_{n-1}\}$ of $\mathcal{R}(f)$.
There must be a vector $v \in \mathcal{F}^n$ such that, $B \cup \{v\}$ is linearly independent, hence is basis of $\mathcal{F}^n$.
But, v does not have a pre-image, as $v \notin \mathcal{R}(f)$. □

Determinant of a Matrix

Working on the Determinant

Consider the matrix,

$$A = \begin{bmatrix} x_1 & x_3 \\ x_2 & x_4 \end{bmatrix}$$

That means, it is a transformation from $\mathbb{R}^2 \to \mathbb{R}^2$ such that,
$(1,0) \to (x_1, x_2)$ and $(0,1) \to (x_3, x_4)$.
So the unit square goes to the parallelogram given by the side vectors (x_1, x_2) and (x_3, x_4).

Working on the Determinant

Consider the matrix,

$$A = \begin{bmatrix} x_1 & x_3 \\ x_2 & x_4 \end{bmatrix}$$

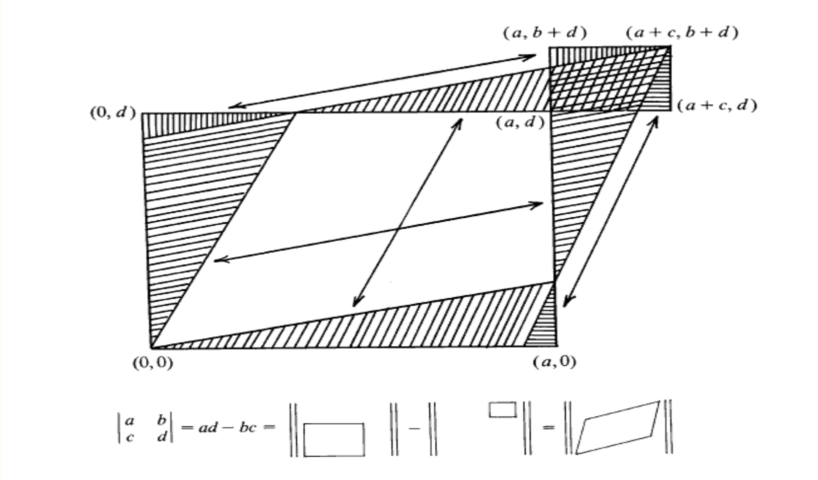
That means, it is a transformation from $\mathbb{R}^2 \to \mathbb{R}^2$ such that, $(1, 0) \to (x_1, x_2)$ and $(0, 1) \to (x_3, x_4)$.
So the unit square goes to the parallelogram given by the side vectors (x_1, x_2) and (x_3, x_4). **Question:** What is the area of the parallelogram?

Determining formula for Determinant

Consider $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ and proof from *Mathematics Magazine, Mar 1985.*

Determining formula for Determinant

Consider $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$ and proof from *Mathematics Magazine, Mar 1985.*



Determinant as Hypervolume

The answer to previous question is $\det(A) = x_1x_4 - x_2x_3$.
So, we generalize a concept called **Determinant**, which is a measure of the Hypervolume of the Hyperparallelopiped, given by the column vectors of A.

Determinant as Hypervolume

The answer to previous question is $\det(A) = x_1x_4 - x_2x_3$.
So, we generalize a concept called **Determinant**, which is a measure of the Hypervolume of the Hyperparallelopiped, given by the column vectors of A.

1. Hence, you cannot think of determinant of non square matrix, since specifying the Hyperparallelopiped uniquely, you require atleast all n columns for $\mathbb{R}^n$.

Determinant as Hypervolume

The answer to previous question is $\det(A) = x_1x_4 - x_2x_3$.
So, we generalize a concept called **Determinant**, which is a measure of the Hypervolume of the Hyperparallelopiped, given by the column vectors of A.

1. Hence, you cannot think of determinant of non square matrix, since specifying the Hyperparallelopiped uniquely, you require atleast all n columns for $\mathbb{R}^n$.
2. If $\det(A) = 0$, then the linear transformation is not invertible. Think that the linear transformation is squeezing many vectors between two of its Hyperspace, at a single Hyperplane, which makes it impossible to retrace back exactly from where these vectors come from.

Properties of Determinant

1. Expanding by row or expanding by column gives same determinant. Because volume is independent of our method of expansion.

Properties of Determinant

1. Expanding by row or expanding by column gives same determinant. Because volume is independent of our method of expansion.
2. If one row or column is 0, then determinant is 0, as $0_{\mathcal{V}}$ is one side of the Hyperparallelopiped.

Properties of Determinant

1. Expanding by row or expanding by column gives same determinant. Because volume is independent of our method of expansion.
2. If one row or column is 0, then determinant is 0, as $0_{\mathcal{V}}$ is one side of the Hyperparallelopiped.
3. Row operation or column operation does not change the volume (extension to the result that area of any parallelogram supported by same parallel lines always remain same).

Properties of Determinant

1. Expanding by row or expanding by column gives same determinant. Because volume is independent of our method of expansion.
2. If one row or column is 0, then determinant is 0, as $0_{\mathcal{V}}$ is one side of the Hyperparallelopiped.
3. Row operation or column operation does not change the volume (extension to the result that area of any parallelogram supported by same parallel lines always remain same).
4. $\det(A^{-1}) = \dfrac{1}{\det(A)}$. Since, if unit cube maps to something of volume $\det(A)$ in range space, an unit cube (or volume) in range space maps back to something of volume $1/\det(A)$, simple unitary method.

Multiplicative Property of Determinant

Theorem

$$\det(AB) = \det(A)\det(B)$$

Multiplicative Property of Determinant

Theorem

$$\det(AB) = \det(A)\det(B)$$

Proof.

Let, f, g be the linear transformations corresponding to A and B. Also, the maps go like $\mathcal{U} \to \mathcal{V} \to \mathcal{W}$.

Multiplicative Property of Determinant

Theorem

$$\det(AB) = \det(A)\det(B)$$

Proof.

Let, f, g be the linear transformations corresponding to A and B. Also, the maps go like $\mathcal{U} \to \mathcal{V} \to \mathcal{W}$.
An unit cube in $\mathcal{U}$ becomes the $\det(A)$ volume of Hyperparallelopiped in $\mathcal{V}$. An unit cube in $\mathcal{V}$ becomes the $\det(B)$ volume of Hyperparallelopiped in $\mathcal{W}$.

Multiplicative Property of Determinant

Theorem

$$\det(AB) = \det(A)\det(B)$$

Proof.

Let, f, g be the linear transformations corresponding to A and B. Also, the maps go like $\mathcal{U} \to \mathcal{V} \to \mathcal{W}$.
An unit cube in $\mathcal{U}$ becomes the $\det(A)$ volume of Hyperparallelopiped in $\mathcal{V}$. An unit cube in $\mathcal{V}$ becomes the $\det(B)$ volume of Hyperparallelopiped in $\mathcal{W}$.
So, an unit cube in $\mathcal{U}$ becomes of volume $\det(A)\det(B)$ in $\mathcal{W}$, when we apply $f \circ g$, whose corresponding matrix is AB. □

Inner Product Space

Inner Product

1. $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathcal{F}$.
2. $\langle x, y \rangle = \langle y, x \rangle$.
3. $\langle ax, y \rangle = a\langle x, y \rangle$.
4. $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$.
5. $\langle x, x \rangle > 0 \quad \forall x \neq 0_{\mathcal{V}}$.

We call $\sqrt{\langle x, x \rangle} = \|x\|$, the norm of x.

Inner Product vs Metric

Inner Product

1. $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathcal{F}$.
2. $\langle x, y \rangle = \langle y, x \rangle$.
3. $\langle ax, y \rangle = a\langle x, y \rangle$.
4. $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$.
5. $\langle x, x \rangle > 0 \quad \forall x \neq 0_{\mathcal{V}}$.

We call $\sqrt{\langle x, x \rangle} = \|x\|$, the norm of x.

Distance function or Metric

1. $d(\cdot, \cdot) : S \times S \to \mathbb{R}$.
2. $d(x, y) = d(y, x)$.
3. $d(x, y) \geq d(x, z) + d(z, y)$.
4. $d(x, x) > 0 \quad \forall x \in S - \{0\}$

Matrix of Inner Product

1. Inner Product looks like a linear operation, hence there could be a matrix for it.
2. Since inner product takes two arguments, the matrix would be multiplied by two vectors to provide its output.

Matrix of Inner Product

1. Inner Product looks like a linear operation, hence there could be a matrix for it.
2. Since inner product takes two arguments, the matrix would be multiplied by two vectors to provide its output.
3. Let, $B = \{v_1, v_2, \dots v_n\}$ be a basis of $\mathcal{V}$.
4. Let, A be a matrix such that,

$$(A)_{ij} = \langle v_i, v_j \rangle$$

Note that, A is symmetric matrix, $A^\top = A$.

Building Inner Product Formula

1. We have, $\mathrm{x} = (x_1, x_2, \ldots x_n)^\top$ and $\mathrm{y} = (y_1, y_2, \ldots y_n)^\top$.
2. Therefore,

$$\begin{aligned}
\langle \mathrm{x}, \mathrm{y} \rangle = \langle \sum_i x_i v_i, \sum_j y_j v_j \rangle &= \sum_i \langle x_i v_i, \sum_j y_j v_j \rangle \\
= \sum_i x_i \langle v_i, \sum_j y_j v_j \rangle &= \sum_i x_i \langle \sum_j y_j v_j, v_i \rangle \\
= \sum_i x_i \sum_j y_j \langle v_j, v_i \rangle &= \sum_i \sum_j x_i y_j \langle v_i, v_j \rangle \\
&= \sum_i \sum_j x_i A_{ij} y_j \\
&= \mathrm{x}^\top A \mathrm{y}
\end{aligned}$$

This A is called the inner product basis matrix. The forms $x^\top A y$ are called **Bilinear** forms.

Formula Inner Product

1. Let, B be the usual n dimensional Euclidean basis. Let, e_i be the vector with all zeros except a one at i-th position.
2. Define, inner product basis matrix to be identity matrix, as usual.
3. Then,

$$\langle x, y \rangle = \sum_i x_i \delta_{ij} y_j$$

where

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Hence, $\langle x, y \rangle = \sum_i x_i y_i$.

Inner Product to Projection

1. Let's say, you want to find the projection (or the component) of x that is aligned with e_1, i.e. what should be the projection of $(x_1, x_2, \dots x_n)$ onto $(1, 0, 0, \dots 0)$?
2. What should be the projection of $(x_1, x_2, \dots x_n)$ onto $(\alpha, 0, 0, \dots 0)$?
3. What should be the projection of $(x_1, x_2, \dots x_n)$ onto $(y_1, y_2, \dots y_n)$?

Inner Product to Projection (Contd.)

1. You want to find projection of x onto y.

Inner Product to Projection (Contd.)

1. You want to find projection of x onto y.
2. This means, you extend this $\{y\}$ to a Orthogonal basis, say, $B = \left\{y, y_1, y_2, \ldots y_{(n-1)}\right\}$, where $y^\top y_i = 0$.

Inner Product to Projection (Contd.)

1. You want to find projection of x onto y.
2. This means, you extend this $\{y\}$ to a Orthogonal basis, say, $B = \left\{y, y_1, y_2, \dots y_{(n-1)}\right\}$, where $y^\top y_i = 0$.
3. Then,

$$x = \alpha y + \alpha_1 y_1 + \alpha_2 y_2 + \dots \alpha_{n-1} y_{n-1}$$

and αy is the part or component of x in the direction of y.

Inner Product to Projection (Contd.)

1. You want to find projection of x onto y.
2. This means, you extend this $\{y\}$ to a Orthogonal basis, say, $B = \left\{y, y_1, y_2, \ldots y_{(n-1)}\right\}$, where $y^\top y_i = 0$.
3. Then,

$$x = \alpha y + \alpha_1 y_1 + \alpha_2 y_2 + \ldots \alpha_{n-1} y_{n-1}$$

   and αy is the part or component of x in the direction of y.
4. So, you consider;

$$\begin{aligned} x &= \alpha y + \alpha_1 y_1 + \alpha_2 y_2 + \ldots \alpha_{n-1} y_{n-1} \\ \Rightarrow \langle x, y\rangle &= \alpha \langle y, y\rangle + \alpha_1 \langle y_1, y\rangle + \cdots + \alpha_{n-1} \langle y_{n-1}, y\rangle \\ \Rightarrow \langle x, y\rangle &= \alpha \langle y, y\rangle \\ \Rightarrow \frac{\langle x, y\rangle}{\langle y, y\rangle} &= \alpha \end{aligned}$$

Answer to the first question

- Now, we go back to the very first question. Why the formula for inner product gives cosine of the angle between the vectors.

Answer to the first question

- Now, we go back to the very first question. Why the formula for inner product gives cosine of the angle between the vectors.
- Note that, geometrically, it is easy to see, the projection of x onto y is $x \cos \theta$.

Answer to the first question

- Now, we go back to the very first question. Why the formula for inner product gives cosine of the angle between the vectors.
- Note that, geometrically, it is easy to see, the projection of x onto y is $x \cos\theta$.
- From the discussion before, the projection was αy, and hence, $x \cos\theta = \alpha y$.

Answer to the first question

- Now, we go back to the very first question. Why the formula for inner product gives cosine of the angle between the vectors.
- Note that, geometrically, it is easy to see, the projection of x onto y is $x \cos\theta$.
- From the discussion before, the projection was αy, and hence, $x \cos\theta = \alpha y$.
- Therefore, taking norm and equating,

$$\|x \cos\theta\| = \|x\| \cos\theta = \frac{\langle x, y\rangle}{\langle y, y\rangle} \|y\|$$

and we end up,

$$\cos\theta = \frac{\langle x, y\rangle}{\|x\| \|y\|}$$

Eigen Values

Eigen Value as Projection

- In a projection of x onto y, we try to find the best part of x which can be made parallel to y.
- Eigenvalues are kind of projection of a linear transformation. You kind of retain best parts of a linear transformation, which explains them most.

Definition

A scalar $\lambda \in \mathcal{F}$ is said to be an eigenvalue with corresponding eigenvector $v \in \mathcal{V}$ of the matrix A if;

$$Av = \lambda v$$

with $v \neq 0$.

Recipe for Eigen values

1. Since, $Av = \lambda v$, we have, $(A - \lambda I)v = 0$.

Recipe for Eigen values

1. Since, $Av = \lambda v$, we have, $(A - \lambda I)v = 0$.
2. That means, $d\left(\mathcal{N}(A - \lambda I)\right) > 0$.
3. That means, by rank nullity theorem, $\rho(A - \lambda I) < n$, if A was $n \times n$ matrix.
4. That means, $(A - \lambda I)$ is not invertible.

Recipe for Eigen values

1. Since, $Av = \lambda v$, we have, $(A - \lambda I)v = 0$.
2. That means, $d\left(\mathcal{N}(A - \lambda I)\right) > 0$.
3. That means, by rank nullity theorem, $\rho(A - \lambda I) < n$, if A was $n \times n$ matrix.
4. That means, $(A - \lambda I)$ is not invertible.
5. So, $\det(A - \lambda I) = 0$. Hence, eigenvalues are roots of the polynomial $\det(A - xI) = 0$.

Recipe for Eigen values

1. Since, $Av = \lambda v$, we have, $(A - \lambda I)v = 0$.
2. That means, $d\left(\mathcal{N}(A - \lambda I)\right) > 0$.
3. That means, by rank nullity theorem, $\rho(A - \lambda I) < n$, if A was $n \times n$ matrix.
4. That means, $(A - \lambda I)$ is not invertible.
5. So, $\det(A - \lambda I) = 0$. Hence, eigenvalues are roots of the polynomial $\det(A - xI) = 0$.

Exercise

Find eigenvalues of a diagonal matrix with entries $a_1, a_2, \ldots a_n$.

Find eigenvalues of the matrix $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$

Decomposition with Eigen values

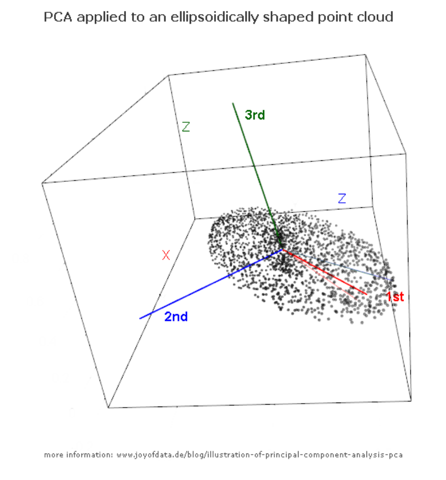
Let, B be a basis with $v_1, v_2, \dots v_n$, each of which is an eigenvector, with corresponding eigenvalue $\lambda_1 > \lambda_2 > \cdots > \lambda_n$.
Then,

$$\begin{aligned} Ax &= A(\sum_i \alpha_i v_i) \\ &= \sum_i \alpha_i A v_i \\ &= \sum_i \alpha_i \lambda_i v_i \\ &\approx \alpha_1\lambda_1 v_1 + \alpha_2\lambda_2 v_2 + \cdots + \alpha_k\lambda_k v_k \end{aligned}$$

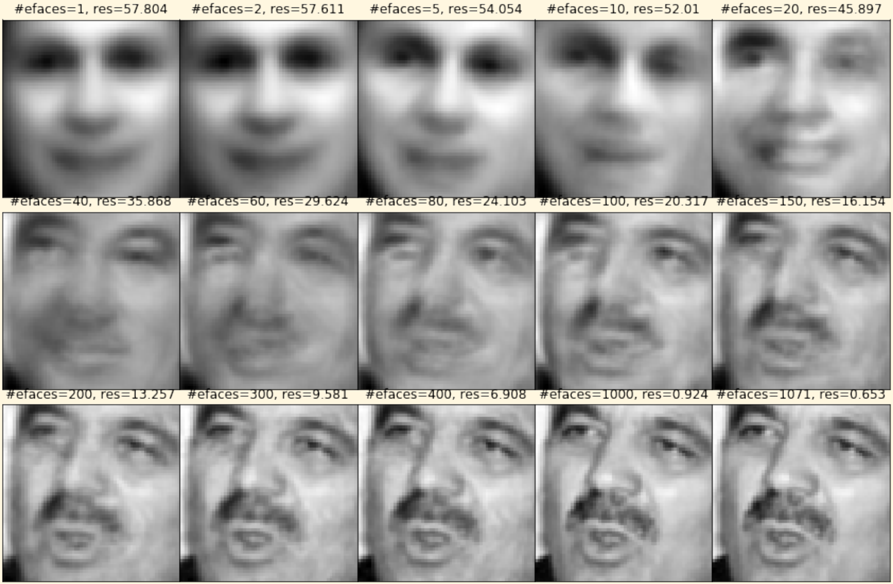
where $k << n$.
Instead of remembering n^2 numbers to specify A, we can simply remember $k(n+1)$, numbers (kn for the eigenvectors and k many for eigenvalues).

Example: Principal Component Analysis

Reference: Math stack exchange.



Example: Eigen Face (from Sandipanweb Wordpress)



STAY HEALTHY & STAY SAFE