Introduction to Linear Algebra
- Gilbert Strang

11

Linear Transformations

ROCK AND ROLL

soorajss1729@gmail.com

+91-9400635788

INDEX

11

NAME: SOORAJ S. STD.: ______ SEC.: ______ ROLL NO.: ______ SUB.: ______

S. No.	Date	Title	Page No.	Teacher's Sign / Remarks
		INTRODUCTION TO		
		LINEAR ALGEBRA		
		– Gilbert Strang, MIT		
		(5th Edition)		

GKC0819079 B

8

LINEAR TRANSFORMATIONS

When a matrix 'A' multiplies a vector 'v', it "transforms" v into another vector, Av.
In goes v, outcomes $T(v) = Av$.

A transformation T follows the same idea as a function. In goes a number x, outcomes $f(x)$.

The deeper goal is to see all vectors v at once. We are transforming the whole space V when we multiply every v by A.

* A transformation T assigns an output $T(v)$ to each input vector v in V.

The transformation is linear if it meets these requirements for all v and w:

ⓐ $T(v+w) = T(v) + T(w)$

ⓑ $T(cv) = c\,T(v)$ for all c.

- If the input is $v=0$, the output must be $T(v)=0$.

$$\boxed{\text{Linear transformation}: \; T(cv+dw) = c\,T(v) + d\,T(w)}$$

Shift is not linear :

$$T(v+w) = v+w+u_0 \neq (v+u_0)+(w+u_0) = T(v)+T(w)$$

Exception is,

Identity transformation, when $u_0 = 0$. :

$T(v) = v$

— linear.

Input space V is the same as the output space W.

The linear-plus-shift transformation $T(v) = Av + u_0$ is called affine.

— not linear.

Ex:1 Choose a fixed vector $a = (1, 3, 4)$ and let $T(v) = a \cdot v$:

The input: $v = (v_1, v_2, v_3)$

The output is : $T(v) = a \cdot v = v_1 + 3v_2 + 4v_3$

$\longrightarrow$ Dot products are linear

The inputs come from 3D space, so $V = \mathbb{R}^3$.
The outputs are just #'s, so the output space is $W = \mathbb{R}^1$.

We are multiplying by the row matrix $A = \begin{bmatrix} 1 & 3 & 4 \end{bmatrix}$ ∴ Then $T(v) = Av$.

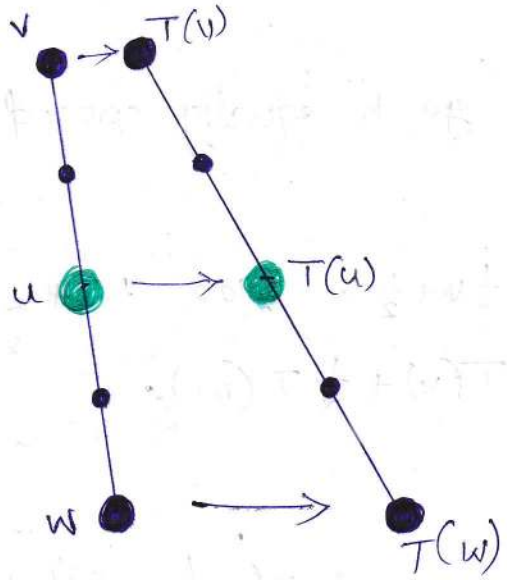
If the o/p involves squares, or products or lengths, v_1^2 or $v_1 v_2$ or $\|v\|$, then T is not linear.

Ex: 2 The length $T(v) = \|v\|$ is not linear.

Ex: 3 (Rotation) T is the transformation that rotates every vector by $30°$. The domain of T is the xy-plane (all i/p vectors v). The range of T is also the xy-plane (all rotated vectors $T(v)$).

Rotation is linear.

□


v
T(v)
u
T(u)
w
T(w)

Linearity : Every point on the input line goes onto the output line.

Equally spaced points go to equally spaced points.

ie., The middle point $u = \frac{1}{2}v + \frac{1}{2}w$ goes to the middle point $T(u) = \frac{1}{2}T(v) + \frac{1}{2}T(w)$.

* Linear transformation keeps straight lines straight

Linearity: $u = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$

transform to

$$T(u) = c_1 T(v_1) + c_2 T(v_2) + \cdots + c_n T(v_n)$$

* If you know $T(v)$ for all vectors $v_1, v_2, \ldots, v_n$ in a basis.

Then you know $T(u)$ for every vector u in the space.

Ex: 4 The transformation T takes the derivative of the input : $T(u) = \frac{du}{dx}$.

$u = 6 - 4x + 3x^2$.

basis vectors : $1, x, x^2$

derivatives : $0, 1, 2x$

Linearity : $\frac{du}{dx} = 6(0) - 4(1) + 3(2x) = -4 + 6x$

• All of calculus depends on linearity!

Precalculus finds a few key derivatives, for x^n and $\sin x$, $\cos x$ and e^x. Then linearity applies to all their combinations.

I would say that the only rule special to calculus is the chain rule. That produces the derivative of a chain of functions $f(g(x))$.

- Nullspace of $T(u) = \frac{du}{dx}$

Solve $T(u) = 0$

The derivative is zero when u is a constant function. So the 1D nullspace is a line in function space - all multiples of the special solution $u = 1$.

Column space of $T(u) = \frac{du}{dx}$

In our example, the i/p space contains all quadratics $a + bx + cx^2$.

The o/p (~~are~~ the column space) are all linear functions $b + 2cx$.

Counting theorem:

$$\dim(\text{column space}) + \dim(\text{nullspace}) = 2 + 1 = 3 = \dim(\text{input space})$$

Matrix for $\frac{d}{dx}$?

linear transformation, $T = \frac{d}{dx}$

$v_1, v_2, v_3 : 1, x, x^2$

$\frac{dv_1}{dx} = 0$, $\underbrace{\frac{dv_2}{dx} = 1}_{= v_1}$, $\frac{dv_3}{dx} = 2x = 2v_2$

The 3D i/p space V (= quadratics) transforms to the 2D o/p space W (= linear functions).

$$\boxed{A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \text{matrix form of the derivative } T = \frac{d}{dx}.}$$

Input u
$a + bx + cx^2$: $Au = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b \\ 2c \end{bmatrix}$. output $\frac{du}{dx} = b + 2cx$

Ex: 5   Integration T^+ is also linear

- The fundamental theorem of Calculus says that integration is the (pseudo) inverse T^+ of differentiation.

I can't say "inverse of T" when the derivative of 1 is 0.

Input v: $D + Ex$

$$A^+ v = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} D \\ E \end{bmatrix} = \begin{bmatrix} 0 \\ D \\ \frac{1}{2}E \end{bmatrix}$$

: output = integral of v

$T^+(v) = Dx + \frac{1}{2}Ex^2$

$$A^+A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$AA^+ = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Ex: 6. Project every 3D vector onto the horizontal plane $z=1$. The vector $v=(x,y,z)$ is transformed to $T(v)=(x,y,1)$. This transformation is not linear.

It doesn't even transform $v=0$ into $T(v)=0$.

Ex: 7. Suppose 'A' is an invertible matrix.

$$T(v+w) = Av + Aw = T(v) + T(w).$$

Another linear transformation is multiplication by A^{-1}. This produces the inverse transformation T^{-1}; which brings every vector $T(v)$ back to v:

$T^{-1}(T(v)) = v$ matches the matrix multiplication

$$A^{-1}(Av) = v$$

If $T(v) = Av$ and $S(u) = Bu$, then the product $T(S(u))$ matches the product ABu.

Are all linear transformations from $V = \mathbb{R}^n$ to $W = \mathbb{R}^m$ produced by matrices ?

When a linear T is described as a rotation or projection or __________,

there is always a matrix 'A' hiding behind T ?

$T(v)$ is always Av.

This is an approach to linear algebra that doesn't start with matrices. We still end up with matrices - after we choose an i/p basis and o/p basis.

Note: Transformations have a language of their own.

Range of T: set of all outputs $T(v)$.
Range corresp. to column space

Kernel of T: set of all inputs for which $T(v) = 0$.
Kernel corresp. to nullspace.

- The range is in the output space W. The kernel is in the input space V.

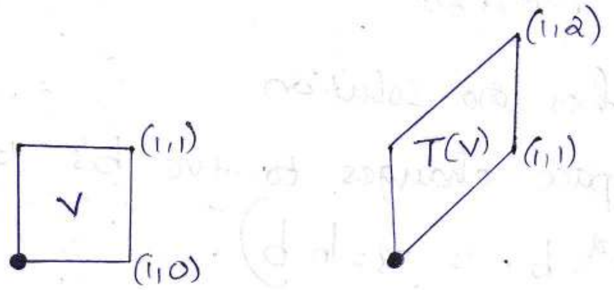
When T is multiplication by a matrix, $T(v) = Av$, range is column space and kernel is nullspace.

$\boxed{8.1(A)}$ The elimination matrix $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ gives a shearing transformation from (x, y) to $T(x, y) = (x, x+y)$. If the inputs fill a square, draw the transformed square.

Ans: The points $(1,0)$ and $(2,0)$ on the x-axis transform by T to $(1,1)$ and $(2,2)$ on the 45° line. Points on the y-axis are not moved: $T(0,y) = (0,y)$: eigenvectors with $\lambda = 1$.

Vertical lines slide up
This is the shearing
squares go to parallelograms

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



$\boxed{8.1(B)}$ A non-linear transformation T is invertible if every b in the output space comes from exactly one x in the input space: $T(x) = b$ always has exactly one solution.

Which of these transformations (on real # x) is invertible and what is T^{-1}?

$T_1(x) = x^2$, $T_2(x) = x^3$, $T_3(x) = x+9$, $T_4(x) = e^x$

$T_5(x) = \frac{1}{x}$ for non zero x's

Ans: $T_1(x) = x^2$ is not invertible

$x^2 = 1$ has 2 solutions & $x^2 = -1$ has no solution

$T_4(x) = e^x$ is not invertible

$e^x = -1$ has no solution

(If the o/p space changes to +ve b's then the inverse of $e^x = b$ is $x = \ln b$).

T_2, T_3, T_5 are invertible. $x^3 = b$, $x+9 = b$ and $\frac{1}{x} = b$ have one solution x.

$x = T_2^{-1}(b) = b^{1/3}$

$x = T_3^{-1}(b) = b - 9$

$x = T_5^{-1}(b) = \frac{1}{b}$

* Let 'A' be a 2×2 matrix. Then the area of the parallelogram generated by the columns of 'A' is $\det(A)$.

* Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation with matrix 'A'. Let R be a region in $\mathbb{R}^2$. Then:

Stack 3/1/2021

$$\boxed{\text{Area}(T(R)) = |\det(A)| \cdot \text{Area}(R)}$$

Note: this theorem will work in higher dimensions too (using volume for $\mathbb{R}^3$ and a higher analog of volume in $\mathbb{R}^n$).

The Matrix of a Linear Transformation

⇒ Assign a matrix to every linear transformation

For ordinary column vectors, the input v is in $V = \mathbb{R}^n$ and the output $T(v)$ is in $W = \mathbb{R}^m$. The matrix 'A' for this transformation will be $m \times n$.
Our choice of bases in V and W will decide A.

The standard basis vectors for $\mathbb{R}^n$ and $\mathbb{R}^m$ are the columns of I. That choice leads to a standard basis.

But these spaces also have other bases, so the same transformation T is represented by other matrices.

→ A main scheme of linear algebra is to choose the bases that give the best matrix (a diagonal matrix) for T.

All vectors spaces V and W have bases. Each choice of those bases leads to a matrix for T. When the i/p basis is different from the o/p basis, the matrix for $T(v) = v$ will not be the identity I. It'll be the "change of basis matrix".

Key Idea

- If we know $T(v)$ for the i/p basis vectors v_1 to v_n.

- Columns 1 to n of the matrix will contain those outputs $T(v_1)$ to $T(v_2)$

- A times c = matrix times vector = combination of those n columns

- Ac is the correct combination

$$c_1 T(v_1) + \cdots\cdots + c_n T(v_n) = T(v)$$

$\Longrightarrow$ Every v is a unique combination $c_1v_1 + \cdots + c_nv_n$ of the basis vectors v_j. Since T is a linear transformation, $T(v)$ must be the same combination $c_1T(v_1) + \cdots + c_nT(v_n)$ of the outputs $T(v_j)$ in the columns.

Ex: 1. Suppose T transforms $v_1 = (1,0)$ to $T(v_1) = (2,3,4)$
Suppose the 2nd basis vector $v_2 = (0,1)$ goes to $T(v_2) = (5,5,5)$. If T is linear from $\mathbb{R}^2$ to $\mathbb{R}^3$ then its standard matrix is 3×2.
Those output $T(v_1)$ and $T(v_2)$ go into the columns of A:

$$A = \begin{bmatrix} 2 & 5 \\ 3 & 5 \\ 4 & 5 \end{bmatrix}$$

$c_1 = 1$ & $c_2 = 1$ give

$$T(v_1 + v_2) = \begin{bmatrix} 2 & 5 \\ 3 & 5 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$

Change of Basis

Suppose the i/p space $V = \mathbb{R}^2$ is also the o/p space $W = \mathbb{R}^2$. Suppose that $T(v) = v$ is the identity transformation.

Expect : matrix is I

But it only happens when the i/p Basis is the same as the o/p Basis.

$T(v) = v$

changing basis from the v's to the w's.

Each v is a combination of w_1 and w_2.

Input Basis : $\begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 3 & 8 \end{bmatrix}$

Output Basis : $\begin{bmatrix} w_1 & w_2 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 1 & 2 \end{bmatrix}$

$v_1 = 1w_1 + 1w_2$

$v_2 = 2w_1 + 3w_2$

We apply the identity transformation T to each input basis vector : $T(v_1) = v_1$ and $T(v_2) = v_2$. Then we write those outputs v_1 and v_2 in the output basis w_1 and w_2.

$$v_1 = 1w_1 + 1w_2$$
$$v_2 = 2w_1 + 3w_2$$

$v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ & $v_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

Matrix B for change of basis : $\begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} w_1 & w_2 \end{bmatrix}\begin{bmatrix} B \end{bmatrix} = \begin{bmatrix} w_1 & w_2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$

$$\boxed{V = WB \longrightarrow B = W^{-1}V}$$

$P_{W \leftarrow V}$

→ When the i/p basis is in the columns of a matrix V, and the o/p basis is in the columns of W, the change of basis matrix for $T = I$ is $\underline{B = W^{-1}V}$.

Think

Suppose the same vector u is written in the input basis of V's and the o/p basis of W's.

$$u = c_1 v_1 + \cdots + c_n v_n$$
$$u = d_1 w_1 + \cdots + d_n w_n$$

$$\Leftrightarrow \begin{array}{c} T(v_1) \quad T(v_n) \\ \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \end{array} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} w_1 & \cdots & w_n \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$$

$$\Leftrightarrow Vc = Wd$$

$$\boxed{d = W^{-1}Vc \Longrightarrow B = W^{-1}V}$$

* The formula $B = W^{-1}V$ produces one of the world's greatest mysteries:

When the standard basis $V = I$ is changed to a different basis W,
the change of basis matrix is not W but W^{-1}.

$\begin{bmatrix} x \\ y \end{bmatrix}$ in the standard basis has coefficients

$\begin{bmatrix} w_1 & w_2 \end{bmatrix}^{-1} \begin{bmatrix} x \\ y \end{bmatrix}$ in the w_1, w_2 basis.

* The change of basis matrix for orthonormal bases is unitary.

Construction of the Matrix

We construct a matrix for any linear transformation.

Suppose,

T transforms the space V (n-dimensional) to the space W (m-dimensional).

We choose a basis $v_1, \ldots, v_n$ for V and we choose a basis $w_1, \ldots, w_n$ for W.

The matrix 'A' will be $m \times n$.

To find the 1st column of A, apply T to the 1st basis vector v_1.

$T(v_1) = a_{11} w_1 + a_{21} w_2 + \cdots + a_{m1} w_m$ of the output basis for W.

→ These numbers $a_{11}, a_{21}, \ldots, a_{m1}$ go into the 1st column of A.

Ex: 3. The i/p basis of V's : $1, x, x^2, x^3$

The o/p basis of W's : $1, x, x^2$

Then,

T takes the derivative : $T(v) = \frac{dv}{dx}$

and A = "derivative matrix"

If $v = c_1 + c_2 x + c_3 x^2 + c_4 x^3$,

then $\frac{dv}{dx} = c_2 + 2c_3 x + 3c_4 x^2$

$$Ac = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix} = \begin{bmatrix} c_2 \\ 2c_3 \\ 3c_4 \end{bmatrix}$$

- The j^{th} column of 'A' is found by applying T to the j^{th} basis vector V_j.

$$T(V_j) = \text{combination of o/p basis vectors}$$
$$= a_{1j} W_1 + a_{2j} W_2 + \cdots + a_{mj} W_m$$

These numbers a_{ij} go into A. The matrix is constructed to get the basis vectors right. Then linearity gets all other vectors right. Every v is a combination $c_1 v_1 + \cdots + c_n v_n$, and $T(v)$ is a combination of the W's.

When 'A' multiplies the vector $c = (c_1, \cdots, c_n)$ in the v combination, Ac produces the coefficients in the T(v) combination. This is because matrix multiplication (combining columns) is linear like T.

The matrix 'A' tells us what 'T' does. Every linear combination from V to W can be converted to a matrix. This matrix depends on the bases.

Ex: 4 For the integral $T^{+}(v)$, the 1st basis function is again 1. Its integral is the 2nd basis function.

The integral of $d_1 + d_2 x + d_3 x^2$

is $d_1 x + \frac{d_2 x^2}{2} + \frac{d_3 x^3}{3}$.

$$A^{+}d = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 0 \\ d_1 \\ \frac{1}{2} d_2 \\ \frac{1}{3} d_3 \end{bmatrix}$$

If you integrate a function and then differentiate, you get back to the start. So $AA^{+} = I$. But, if you differentiate before integrating, the constant term is lost. So $A^{+}A$ is not I.

$$T^{+}T(1) = 0$$

This matches $A^{+}A$, whose 1st column is all zero.

The derivative T has a kernel (the constant functions). Its matrix 'A' has a nullspace.

⇒ ① Linear transformations T are everywhere - in calculus, and differential equations, and linear algebra.

② Spaces other than $\mathbb{R}^n$ are important - we had functions in V and W.

③ If we differentiate and then integrate, we can multiply their matrices A^+A.

Matrix products AB Match Transformations TS

Two linear transformations T & S are represented by 2 matrices A and B.

When we apply the transformation T to the output from S, we get TS by this rule:

$$(TS)(u) = T(S(u)).$$

The output $S(u)$ becomes the input to T.

When we multiply the matrix A to the output from B, we multiply AB by this rule:

$$(AB)(x) = A(Bx).$$

The output Bx becomes the input to A.

$\Rightarrow$ Matrix multiplication gives the correct matrix AB to represent TS.

The transformation S is from a space U to V. Its matrix B uses a basis $u_1, \dots, u_p$ for U and a basis $v_1, \dots, v_n$ for V. That matrix is $n \times p$. The transformation T is from V to W as before. Its matrix A must use the same basis $v_1, \dots, v_n$ for V — this is the o/p space for S and the i/p space for T. Then the matrix AB matches TS.

The linear transformation TS starts with any vector u in U, goes to $S(u)$ in V and then to $T(S(u))$ in W. The matrix AB starts with any x in $\mathbb{R}^p$, goes to Bx in $\mathbb{R}^n$, and then to ABx in $\mathbb{R}^m$. The matrix AB correctly represents TS :

$$TS : U \to V \to W$$

$$AB : (m \text{ by } n)(n \text{ by } p) = (m \text{ by } p)$$

The output $T(S(u))$ matches the output ABx.

Product of transformations TS matches product of matrices AB.

Ex: 5 S rotates the plane by θ and T also rotates by θ. Then TS rotates by 2θ. This transformation. T^2 corresp. to the rotation matrix A^2 through 2θ.

$T = S \qquad A = B \qquad T^2 =$ rotation by 2θ

$$A^2 = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix} \quad \text{—— (4)}$$

By matching $(\text{transformation})^2$ with $(\text{matrix})^2$, we pick up the formulas for $\cos 2\theta$ and $\sin 2\theta$.

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta - \sin^2\theta & -2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{bmatrix} \quad \text{—— (5)}$$

Comparing (4) and (5) produces $\cos 2\theta = \cos^2\theta - \sin^2\theta$ and $\sin 2\theta = 2\sin\theta\cos\theta$.

— Trigonometry (the double angle rule) comes from linear algebra.

Choosing the Best Bases

- Choose bases that diagonalize the matrix.

Eigenvectors: If T transforms $\mathbb{R}^n$ to $\mathbb{R}^n$, its matrix A is square. But using the standard basis, that matrix A is probably not diagonal. If there are n independent eigenvectors, choose those as the input and output basis.
In this good basis the matrix for T is the diagonal eigenvalue matrix Λ.

Ex: 7 The projection matrix T projects every $v = (x, y)$ in $\mathbb{R}^2$ onto the line $y = -x$.

Using the standard basis, $v_1 = (1, 0)$ projects to $T(v_1) = (\frac{1}{2}, -\frac{1}{2})$. For $v_2 = (0, 1)$ the projection is $T(v_2) = (-\frac{1}{2}, \frac{1}{2})$.

Projection Matrix
standard bases
Not diagonal : $A = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$ has $A^T = A$ and $A^2 = A$

When the basis vectors are eigenvectors, the matrix becomes diagonal.

$v_1 = w_1 = (1, -1)$ projects to itself : $T(v_1) = v_1$ and $\lambda_1 = 1$

$v_2 = w_2 = (1, 1)$ projects to zero : $T(v_2) = 0$ and $\lambda_2 = 0$

Eigenvector bases
Diagonal Matrix : The new matrix is

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \Lambda$$

Eigen vectors are perfect basis vectors: They produce the eigenvalue matrix Λ.

Let $T: \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation & we want to find the matrix defined by this linear transformation :

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \sum_{i=1}^{n} x_i \hat{e}_i$$

Since T is linear,

$$T(\vec{x}) = \sum_{i=1}^{n} x_i \, T(\hat{e}_i)$$

$$= \begin{bmatrix} | & | & & | \\ T(\hat{e}_1) & T(\hat{e}_2) & \cdots & T(\hat{e}_n) \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A\vec{x}$$

What about other choices of input basis = output basis ?

The new matrix for T is similar to A.

$$\boxed{A_{new} = B^{-1}AB}$$ in the new basis of b's is similar to A in the standard basis.

A_{vw} $\quad C^{-1}B$

$$T_{new} = A_{new} = \left[T_{new}(b_1), \dots\dots, T_{new}(b_n)\right]$$

section change of basis. (1st define $T(v)$ then write it in terms of $w_1, w_2 \dots$)

$$= \left[\cancel{A(v_1)}(Ab_1)_{new} \quad \dots \quad (Ab_n)_{new}\right]$$

$$= \left[B^{-1}Ab_1, \dots\dots B^{-1}Ab_n\right]$$

$$= B^{-1}A\left[b_1 \dots b_n\right] = B^{-1}AB$$

$$[T]_B = P_{B\leftarrow E}\,[T]\,P^{-1}_{B\leftarrow E}$$

where, $P_{B\leftarrow E} = B^{-1}$

Assume that V is some vector space and $\dim V = n < \infty$.

Let $B = \{b_1, b_2, \cdots, b_n\}$ and $C = \{c_1, c_2, \cdots, c_n\}$ be a bases of V.

For any vector $\vec{v} \in V$, let $[\vec{v}]_B$ and $[\vec{v}]_C$ be its coordinate vectors w.r.t the bases B & C, respectively. These vectors are related by the formula,

$$\boxed{[\vec{v}]_C = P_{C \leftarrow B} [\vec{v}]_B}$$

where, $P_{C \leftarrow B}$ is the change of coordinates matrix from B to C, given by

$$\boxed{P_{C \leftarrow B} = \left[[\vec{b}_1]_C \ [\vec{b}_2]_C \cdots\cdots [\vec{b}_n]_C \right]} \boxed{= C^{-1} B}$$

If $T : V \to V$ is a linear transformation, then its matrix in the basis B is given by

$$\boxed{[T]_B = \left[[T(b_1)]_B \ [T(b_2)]_B \cdots\cdots [T(b_n)]_B \right]}$$

$$\boxed{[T(\vec{v})]_B = [T]_B [\vec{v}]_B \text{ for all } \vec{v} \in V}$$

The matrix of T in the basis B and its matrix in the basis C are related by the formula,

$$\boxed{[T]_C = P_{C\leftarrow B} [T]_B P_{C\leftarrow B}^{-1}}$$

* The matrices of T in 2 different bases are similar..

$$P_{C\leftarrow B} = C^{-1}B \Longrightarrow \boxed{P_{C\leftarrow B}^{-1} = B^{-1}C = P_{B\leftarrow C}}$$

$$\boxed{P_{B\leftarrow C}\, P_{C\leftarrow D} = B^{-1}C\,C^{-1}D = B^{-1}D = P_{B\leftarrow D}}$$

Let V and W be non-trivial vector spaces, with $\dim V = n$ and $\dim W = m$.

Let $B = \{b_1, b_2, \ldots, b_n\}$ and $C = \{w_1, w_2, \ldots, w_m\}$ be ordered bases for V and W, respectively.

Let $T : V \to W$ be a linear transformation, then the matrix of T relative to the bases B and C is given by,

$$\boxed{[T]_{BC} = \Big[[T(b_1)]_C \;\; [T(b_2)]_C \cdots\cdots [T(b_n)]_C \Big]}$$

$$\boxed{[T]_{BC} [v]_B = [T(v)]_C}, \text{ for all } v \in V$$

Let V & W be 2 non-trivial finite dimensional vector spaces with ordered Bases B & C, respectively. Let $T: V \to W$ be a linear transformation with matrix A_{BC} w.r.t the Bases B & C.

Suppose that D and G are other ordered Bases for V and W, respectively.

$$\boxed{[T]_{DG} = P_{G \leftarrow C} [T]_{BC} P^{-1}_{D \leftarrow B}}$$

$$\boxed{[T]_{DG} = P_{G \leftarrow B} [T]_{BB} P^{-1}_{D \leftarrow B}}$$

Proof

$$\left[T\right]_{DG} = \left[\left[T(d_1)\right]_{G} \cdots\cdots \left[T(d_n)\right]_{G} \right]$$

$$= \left[P_{G\leftarrow E} \left[T(d_1)\right] \cdots\cdots P_{G\leftarrow E}\left[T(d_n)\right] \right]$$

$$= P_{G\leftarrow E} \left[\left[T\right] d_1 \cdots\cdots \left[T\right] d_n \right]$$

$$= P_{G\leftarrow E} \left[T\right] \left[d_1 \cdots d_n \right]$$

$$= P_{G\leftarrow E} \left[T\right] P^{-1}_{D\leftarrow E}$$

$$\left[T\right]_{BC} = \left[\left[T(b_1)\right]_{C} \cdots\cdots \left[T(b_n)\right]_{C} \right]$$

$$= \left[P_{C\leftarrow E} \left[T(b_1)\right] \cdots\cdots P_{C\leftarrow E}\left[T(b_n)\right] \right]$$

$$= P_{C\leftarrow E} \left[\left[T\right] b_1 \cdots\cdots \left[T\right] b_n \right]$$

$$= P_{C\leftarrow E} \left[T\right] \left[b_1 \cdots b_n \right]$$

$$= P_{C\leftarrow E}\,[T]\,P^{-1}_{B\leftarrow E}$$

$$[T]_{DG} = P_{G\leftarrow E}\,[T]\,P^{-1}_{D\leftarrow E}$$

$$= P_{G\leftarrow E}\,P^{-1}_{C\leftarrow E}\,\underbrace{P_{C\leftarrow E}\,[T]\,P^{-1}_{B\leftarrow E}}\,P_{B\leftarrow E}\,P^{-1}_{D\leftarrow E}$$

$$= P_{G\leftarrow E}\,P^{-1}_{C\leftarrow E}\cdot[T]_{BC}\,P_{B\leftarrow E}\,P^{-1}_{D\leftarrow E}$$

$$= P_{G\leftarrow E}\,P_{E\leftarrow C}\,[T]_{BC}\,P_{B\leftarrow E}\,P_{E\leftarrow D}$$

$$= P_{G\leftarrow C}\,[T]_{BC}\,P_{B\leftarrow D}$$

$$= P_{G\leftarrow C}\,[T]_{BC}\,P^{-1}_{D\leftarrow B}$$

Different spaces V and W, and different bases v's and w's.

When we know T and we choose bases, we get a matrix A.

We can always choose v's and w's that produce a diagonal matrix. This will be the singular value matrix $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_n)$ in the decomposition $A = U\Sigma V^T$.

Singular vectors: The SVD says that $U^{-1}AV = \Sigma$.

The right singular vectors $v_1, \dots, v_n$ will be the i/p basis. The left singular vectors $u_1, \dots, u_m$ will be the output basis.

By the rule for matrix multiplication, the matrix for the same transformation in these new bases is:

$$B_{out}^{-1} A B_{in} = U^{-1} A V = \Sigma$$

* Even when Q is rectangular (matrix with orthonormal columns)

$$Q^T Q = I$$

Q^T is only an inverse from the left.

$$(Qu)\cdot(Qv) = (Qu)^T(Qv) = u^T Q^T Q v$$

$$= u^T v = u \cdot v$$

$\Longrightarrow$ ~~M~~ preserves the length of vectors.

I can't say Σ is "similar" to A.
We are working now with 2 bases, input and output.

$\Longrightarrow$ Σ is "isometric" to A

* $C = Q_1^{-1} A Q_2$ is isometric to A if Q_1 and Q_2 are orthogonal.

Ex.8 To construct the matrix A for the transformation $T = \frac{d}{dx}$, we chose the i/p Basis $1, x, x^2, x^3$ and the o/p basis $1, x, x^2$. The matrix A was simple but unfortunately it wasn't diagonal.

But, we can take each basis in opposite order.

Ans:

Now,
the i/p Basis is $x^3, x^2, x, 1$ and the output Basis is $x^2, x, 1$. The change of Basis matrices B_{in} and B_{out} are permutations.

The matrix for $T(u) = \frac{du}{dx}$ with new bases is the diagonal singular value matrix $B_{out}^{-1} A B_{in} = \Sigma$ with $\sigma's = 3, 2, 1$.

$$B_{out}^{-1} A B_{in} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

8.2(A) The space of 2x2 matrices has these 4 vectors as a Basis.

ILA(15)

8.1(16)

&

ILA (14) Backside

isomorphism

$$V_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ V_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ V_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},\ V_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

T is a linear transformation that transposes every 2x2 matrix. What is the matrix A that represents T in this Basis (o/p basis = i/p basis)?

What is the inverse matrix A^{-1}?

What's the transformation T^{-1} that inverts the transpose operation?

Ans:

$$\left.\begin{aligned} T(V_1) &= V_1 \\ T(V_2) &= V_3 \\ T(V_3) &= V_2 \\ T(V_4) &= V_4 \end{aligned}\right\} \text{gives the 4 columns of } A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} V_1 & V_3 \\ V_2 & V_4 \end{bmatrix}$$

- The space of 2×2 matrices is 4-dimensional. So, the matrix A (for the transpose T) is 4×4.

The nullspace of A is Z and the kernel of T is the zero matrix – the only matrix that transposes to zero.

The eigenvalues of A are $1, 1, 1, -1$.

The inverse matrix A^{-1} is the same as A. The inverse transformation T^{-1} is the same as T^{-1}. If we transpose and transpose again, the final matrix equals the original matrix.

Ex:- Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear operator given by $T([x_1, x_2, x_3]) = [3x_1 + x_2, x_1 + x_3, x_1 - x_3]$. Find the matrix for T w.r.t the standard basis for $\mathbb{R}^3$.

Ans: ~~The standard b~~

$$T(e_1) = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad T(e_2) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad T(e_3) = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

The matrix for T w.r.t the standard basis is

$$[T] = \begin{bmatrix} T(e_1) & T(e_2) & T(e_3) \end{bmatrix} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}$$

Ex:- Let $T : P_3 \to \mathbb{R}^3$ be the linear transformation given by, $T(ax^3 + bx^2 + cx + d) = [4a - b + 3c + 3d, a + 3b - c + 5d, -2a - 7b + 5c - d]$. Find the matrix for T w.r.t the standard bases $B = \{x^3, x^2, x, 1\}$ for P_3 and $C = \{e_1, e_2, e_3\}$ for $\mathbb{R}^3$.

Ans:

$$[T(x^3)]_C = \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix}, [T(x^2)]_C = \begin{bmatrix} -1 \\ 3 \\ -7 \end{bmatrix}, [T(x)]_C = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, [T(1)]_C = \begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}$$

The matrix of T w.r.t the Bases B & C is:

$$[T]_{BC} = \left[[T(x^3)]_C \; [T(x^2)]_C \; [T(x)]_C \; [T(1)]_C \right]$$

$$= \begin{bmatrix} 4 & -1 & 3 & 3 \\ 1 & 3 & -1 & 5 \\ -2 & -7 & 5 & -1 \end{bmatrix}.$$

Ex:- Let $T: P_3 \to P_2$ be the linear transformation given by $T(p) = p'$, where $p \in P_3$. Find the matrix for T w.r.t the standard Bases for P_3 and P_2. Use this matrix to calculate $T(4x^3 - 5x^2 + 6x - 7)$ by matrix multiplication.

Ans: The standard basis for P_3 is $B = \{x^3, x^2, x, 1\}$ & the standard basis for P_2 is $C = \{x^2, x, 1\}$

$T(x^3) = 3x^2$, $T(x^2) = 2x$, $T(x) = 1$, $T(1) = 0$.

$$3x^2 \to \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \quad 2x \to \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \quad 1 \to \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad 0 \to \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$A_{BC} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\left[4x^3 - 5x^2 + 6x - 7\right]_B = \begin{bmatrix} 4 \\ -5 \\ 6 \\ -7 \end{bmatrix}$$

$$\left[T(4x^3 - 5x^2 + 6x - 7)\right]_C = A_{BC} \left[4x^3 - 5x^2 + 6x - 7\right]_B$$

$$= \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ -5 \\ 6 \\ -7 \end{bmatrix} = \begin{bmatrix} 12 \\ -10 \\ 6 \end{bmatrix}$$

Converting back from C-coordinates to polynomials gives.

$$T(4x^3 - 5x^2 + 6x - 7) = 12x^2 - 10x + 6$$

Ex:- Let $T: \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation given by $T([x_1, x_2, x_3]) = [-2x_1 + 3x_3, x_1 + 2x_2 - x_3]$. Find the matrix for T w.r.t the ordered basis $B = \{[1,-3,2], [-4,13,-3], [2,-3,20]\}$ for $\mathbb{R}^3$ and $C = \{[-2,-1], [5,3]\}$ for $\mathbb{R}^2$.

Ans:

$$[T]_{BC} = \Big[[T(b_1)]_C \quad [T(b_2)]_C \quad [T(b_3)]_C \Big]$$

$$T(b_1) = [4 \ \ -7], \quad T(b_2) = [-1, 25], \quad T(b_3) = [56, -24]$$

$$[T(b_1)]_C = P_{C \leftarrow E} [T(b_1)]_E = W^{-1} [T(b_1)]$$

$$W = \begin{bmatrix} -2 & 5 \\ -1 & 3 \end{bmatrix} \longrightarrow W^{-1} = -1 \begin{bmatrix} 3 & -5 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} -3 & 5 \\ -1 & 2 \end{bmatrix}$$

$$[T(b_1)]_C = W^{-1} \{T(b_1)\} = \begin{bmatrix} -3 & 5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ -7 \end{bmatrix} = \begin{bmatrix} -47 \\ -18 \end{bmatrix}$$

$$[T(b_2)]_C = W^{-1} [T(b_2)] = \begin{bmatrix} -3 & 5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 25 \end{bmatrix} = \begin{bmatrix} 128 \\ 51 \end{bmatrix}$$

$$[T(b_3)]_C = W^{-1} [T(b_3)] = \begin{bmatrix} -3 & 5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 56 \\ -24 \end{bmatrix} = \begin{bmatrix} -288 \\ -104 \end{bmatrix}$$

The matrix of T w.r.t the bases B & C is:

$$[T]_{BC} = \begin{bmatrix} -47 & 128 & -288 \\ -18 & 51 & -104 \end{bmatrix}$$

Ex:- Let $T: \mathbb{R}^3 \to \mathbb{R}^3$ be the linear operator given by

$$T[(a,b,c)] = [-2a+b,\ -b-c,\ a+3c]$$

(a) Find the matrix A_{BB} for T w.r.t the standard basis $B = \{e_1 = [1,0,0], e_2 = [0,1,0], e_3 = [0,0,1]\}$ for $\mathbb{R}^3$

(b) Use (a) to find A_{DE} w.r.t the standard bases $D = \{[15,-6,4], [2,0,1], [3,-1,1]\}$ and $E = \{[1,-3,1], [0,3,-1], [2,-2,1]\}$.

Ans:

(a) $T(e_1) = [-2,0,1]$, $T(e_2) = [1,-1,0]$, $T(e_3) = [0,-1,3]$

$$A_{BB} = [T] = \begin{bmatrix} T(e_1) & T(e_2) & T(e_3) \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 1 & 0 \\ 0 & -1 & -1 \\ 1 & 0 & 3 \end{bmatrix}$$

(b) $A_{DE} = P_{E \leftarrow B}\ A_{BB}\ P^{-1}_{D \leftarrow B}$

$$P_{E\leftarrow B} = E^{-1} = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 3 & -2 \\ 1 & -1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -2 & -6 \\ 1 & -1 & -4 \\ 0 & 1 & 3 \end{bmatrix}$$

$$P^{-1}_{D\leftarrow B} = D = \begin{bmatrix} 15 & 2 & 3 \\ -6 & 0 & -1 \\ 4 & 1 & 1 \end{bmatrix}$$

$$A_{DE} = P_{E\leftarrow B}\, A_{BB}\, P^{-1}_{D\leftarrow B}$$

$$= \begin{bmatrix} 1 & -2 & -6 \\ 1 & -1 & -4 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 \\ 0 & -1 & -1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} 15 & 2 & 3 \\ -6 & 0 & -1 \\ 4 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -202 & -32 & -43 \\ -146 & -23 & -31 \\ 83 & 14 & 18 \end{bmatrix}$$

Ex:- Let $T : P_3 \to \mathbb{R}^3$ be the linear transformation given by $T(ax^3 + bx^2 + cx + d) = [c+d, 2b, a-d]$

(a) Find the matrix A_{BC} for T w.r.t the standard bases B (for P_3) and C (for $\mathbb{R}^3$)

(b) Use (a) to find the matrix A_{DE} for T w.r.t the standard bases $B = \{x^3, x^2, x, 1\}$ for P_3 and $C = \{e_1 = [1,0,0], e_2 = [0,1,0], e_3 = [0,0,1]\}$ for $\mathbb{R}^3$

$D = \{x^3 + x^2, x^2 + x, x+1, 1\}$ for P_3 and

$E = \{[-2,1,-3], [1,-3,0], [3,-6,2]\}$ for $\mathbb{R}^3$.

Ans:

(a) $T(x^3) = [0,0,1]$, $T(x^2) = [0,2,0]$, $T(x) = [1,0,0]$

$T(1) = [1,0,-1]$

$$\left[T(ax^3+bx^2+cx+d)\right] = \left[T(x^3) \quad T(x^2) \quad T(x) \quad T(1)\right]$$

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

(b) $A_{DE} = P_{E\leftarrow C} \; A_{BC} \; P^{-1}_{D\leftarrow B}$

$$P_{E\leftarrow C} = E^{-1} = \begin{bmatrix} -2 & 1 & 3 \\ 1 & -3 & -6 \\ -3 & 0 & 2 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} -6 & -2 & 3 \\ 16 & 5 & -9 \\ -9 & -3 & 5 \end{bmatrix}$$

$$P^{-1}_{D\leftarrow B} = P_{B\leftarrow D} = B^{-1} D \; ?$$

To compute $P^{-1}_{D\leftarrow B}$, we need to convert the polynomials in D into vectors in $\mathbb{R}^4$;

$$ax^3+bx^2+cx+d \in P_3 \longrightarrow$$

$(x^3+x^2) \rightarrow [1,1,0,0]$; $(x^2+x) \rightarrow [0,1,1,0]$;

$(x+1) \rightarrow [0,0,1,1]$; $(1) \rightarrow [0,0,0,1]$

$$P^{-1}_{D\leftarrow B} = P_{B\leftarrow D} = D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$A_{DE} = P_{E\leftarrow C}\, A_{BC}\, P^{-1}_{D\leftarrow B}$$

$$= \begin{bmatrix} -6 & -2 & 3 \\ 16 & 5 & -9 \\ -9 & -3 & 5 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & -10 & -15 & -9 \\ 1 & 26 & 41 & 25 \\ -1 & -15 & -23 & -14 \end{bmatrix}$$

The Search for a Good Basis

Pure Algebra: If A is the matrix for a transformation T in the standard basis, then

$B_{out}^{-1} A B_{in}$ is the matrix in the new bases.

The standard basis vectors are the columns of the identity: $B_{in} = I_{n\times n}$ and $B_{out} = I_{m\times m}$

Now, we are choosing special bases to make the matrix clearer and simpler than A.
When $B_{in} = B_{out} = B$, the square matrix $B^{-1}AB$ is similar to A.

Applied Algebra: Applications are all about choosing good bases. Here are 4 important choices for vectors and 3 choices for functions.

① $B_{in} = B_{out} =$ eigenvector matrix X

$$\Longrightarrow X^{-1}AX = \text{eigenvalues in } \Lambda$$

This choice requires 'A' to be a square matrix with 'n' independent eigenvectors.
"A must be diagonalizable".

② $B_{in} = V$ and $B_{out} = U$: singular vectors of A.

$$\Longrightarrow U^{-1}AV = \text{diagonal } \Sigma$$

Σ is the singular value matrix (with $\sigma_1, \dots, \sigma_r$ on its diagonal) when B_{in} and B_{out} are the singular vector matrices V and U

Columns of B_{in} and B_{out} are orthonormal eigenvectors of A^TA and AA^T.

③ $B_{in} = B_{out}$ = generalized eigenvectors of A.

$$\longrightarrow B^{-1}AB = \text{Jordan form, } J$$

'A' is a square matrix but it may only have s independent eigenvectors.

(If $s=n$, then B is X and J is Λ)

In all cases, Jordan constructed $(n-s)$ additional "generalized" eigenvectors, aiming to make the Jordan form J as diagonal as possible:

(i) There are s square blocks along the diagonal of J.

(ii) Each block has one eigenvalue λ, one eigenvector, and 1's above the diagonal.

The good case has n, 1×1 blocks, each containing an eigenvalue. Then J is Λ (diagonal).

Ex:1 This Jordan matrix J has eigenvalues $\lambda = 2, 2, 3, 3$ (2 double eigenvalues). Those eigenvalues lie along the diagonal because J is triangular.

There are 2 independent eigenvectors for $\lambda = 2$, but there is only one line of eigenvectors for $\lambda = 3$.
This will be true for every matrix $C = BJB^{-1}$ that is similar to J.

$$\text{Jordan matrix, } J = \begin{bmatrix} 2 & & & \\ & 2 & & \\ & & \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \end{bmatrix}$$

two 1x1 blocks
One 2x2 block
3 eigenvectors
Eigenvalues 2, 2, 3, 3.

Two eigenvectors for $\lambda = 2$ are $x_1 = (1,0,0,0)$ and $x_2 = (0,1,0,0)$. One eigenvector for $\lambda = 3$ is $x_3 = (0,0,1,0)$.

The "generalized eigenvector" for this Jordan matrix is the 4^{th} standard basis vector $x_4 = (0,0,0,1)$.

The eigenvectors for J (normal & generalized) are just the columns x_1, x_2, x_3, x_4 of the identity matrix.

$(J - 3I) x_4 = x_3$.

The generalized eigenvector x_4 connects to the true eigenvector x_3.

A true x_4 would have $(J-3I)x_4 = 0$, but that doesn't happen here.

Every matrix $C = BJB^{-1}$ that is similar to this J will have true eigenvectors b_1, b_2, b_3 in the 1st 3 columns of B. The 4th column of B will be a generalized eigenvector b_4 of C, tied to the true b_3.

$Bx_3 = b_3$, $Bx_4 = b_4$

$(C - 3I) b_4 = b_3$.

$$(BJB^{-1} - 3I) b_4 = BJx_4 - 3Bx_4 = B(J-3I)x_4$$
$$= Bx_3 = b_3$$

Jordan's theorem $\longrightarrow$ Every square matrix A has a complete set of eigenvectors and generalized eigenvectors. When those go into the columns of B, the matrix $B^{-1}AB = J$ is in Jordan form.

The Jordan Form

For every 'A' we want to choose B so that $B^{-1}AB$ is as nearly diagonal as possible.
When 'A' has a full set of n eigenvectors, they go into the columns of B. Then $B = X$.
The matrix $X^{-1}AX$ is diagonal.

This is the Jordan form of A - when A can be diagonalized.

Suppose, 'A' has s independent eigenvectors. Then it is similar to a Jordan matrix with s blocks. Each block has an eigenvalue on the diagonal with 1's just above it. This block accounts for exactly one eigenvector of A. Then B contains generalized eigenvectors as well as ordinary eigenvectors.

When there are n eigenvectors, all n blocks will be 1×1. In that case $J = \Lambda$.

If 'A' has 's' independent eigenvectors, it is similar to a matrix J that has 's' Jordan blocks $J_1, \ldots, J_s$ on its diagonal. Some invertible matrix B puts A into Jordan form :

Jordan form / Jordan normal form / Jordan canonical form : $B^{-1}AB = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix} = J$

$= J_1 \oplus \cdots \oplus J_s$

Each block J_i has one eigenvalue λ_i, one eigenvector, and 1's just above the diagonal:

Jordan block : $J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$

* Matrices are similar if they share the same Jordan form J, but not otherwise.

* The matrix J is unique upto a permutation of the blocks $J_i(\lambda_i)$.

~~An $n \times n$ matrix 'A' is similar to a matrix in Jordan form iff $\mathbb{C}^n$ has a basis~~

- A canonical form (or) standard form of a mathematical object is a standard way of presenting that object as a mathematical expression. Often, it is one which provides the simplest representation of an object and which allows it to be identified in a unique way.

Ex:- • Jordan normal form is a canonical form for matrix similarity.

• The eqn. of a circle: $(x-h)^2 + (y-k)^2 = r^2$

* An $n \times n$ matrix A is diagonalizable over $\mathbb{C}$ if and only if $\mathbb{C}^n$ has a basis consisting entirely of eigenvectors of 'A'.

The matrix B with these basis vectors as its columns gives a similarity transformation of A with a diagonal matrix whose entries are the eigenvalues of A.
The Jordan form of a diagonalizable matrix consists entirely of 1×1 blocks.

* An $n \times n$ matrix 'A' is similar to a matrix in Jordan form iff $\mathbb{C}^n$ has a basis which can be partitioned into a collection of "strings" of vectors, (generalized eigenvectors)
ie., a typical string consisting of vectors $x_1, x_2, \ldots, x_s$ for which

either $Ax_i = \lambda_i x_i$ (or) $Ax_i = \lambda_i x_i + x_{i-1}$

for each $i = 1, 2, \ldots, s$.

Each string corresponds to an $m \times m$ Jordan block involving the eigenvalue λ_i of 'A', and the set of strings is in one-one correspondence with the set of blocks making up the Jordan form.

→ To prove that 'A' is similar to a matrix in Jordan form it suffices to produce a basis of the above kind (above string of basic vectors).

Ex:-

$$J = \begin{bmatrix} 8 & 1 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 8 & 1 \\ 0 & 8 \end{bmatrix} & & \\ & \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} & \\ & & \begin{bmatrix} 0 \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} J_1 & & \\ & J_2 & \\ & & J_3 \end{bmatrix}$$

$$B^{-1}AB = J \longrightarrow AB = BJ$$

$$A \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{bmatrix} \begin{bmatrix} 8 & 1 & & & \\ 0 & 8 & & & \\ & & 0 & 1 & \\ & & 0 & 0 & \\ & & & & 0 \end{bmatrix}$$

Theorem: Let 'A' be an $n \times n$ complex matrix. Then there exists an invertible matrix B such that;

$$B^{-1} A B = J$$

where J is a Jordan form matrix having the eigenvalues of 'A'.

$$B^{-1} A B = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_s \end{bmatrix} = J_1 \oplus \cdots \oplus J_s = J$$

where, $A_{n \times n}$, $B_{n \times n}$, $J_{n \times n}$

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}, \text{ for } i = 1, 2, \ldots, s$$

total s eigenvector for A.

This Jordan canonical form is unique, except for the order of the Jordan blocks of which it is composed.

Equivalently,

the columns of B consists of a set of independent vectors (generalized eigenvectors) $x_1, x_2, \ldots, x_n$ such that

$$A x_i = \lambda_i x_i \quad \text{(or)} \quad A x_i = \lambda_i x_i + x_{i-1}$$

For every $n \times n$ matrix 'A' with complex entries is similar to a matrix in Jordan canonical form. For each eigenvalue λ of A, there will be one or more sequences of generalized eigenvectors $x_1, \ldots, x_k$ (one sequence for each Jordan block):

$$(A - \lambda I) x_1 = 0$$

$$A x_2 = \lambda x_2 + x_1 \longrightarrow (A - \lambda I)^2 x_2 = 0$$

$$\vdots$$

$$A x_k = \lambda x_k + x_{k-1} \longrightarrow (A - \lambda I)^k x_k = 0$$

The sequence $x_1, x_2, \ldots, x_k$ is called a string (of generalized eigenvectors) headed by the eigenvector x_1.

Proof - Fillipov's inductive proof

When $n=1$, the Jordan canonical form of the matrix $[a]$ is $[a]$ itself.

Assume the existence of a Jordan canonical form for all $r \times r$ matrices, $r = 1, 2, \ldots, n-1$.

ie.,

$A_{r\times r}$ is similar to a Jordan matrix.

~~Assume that $A_{n\times n}$ is singular~~

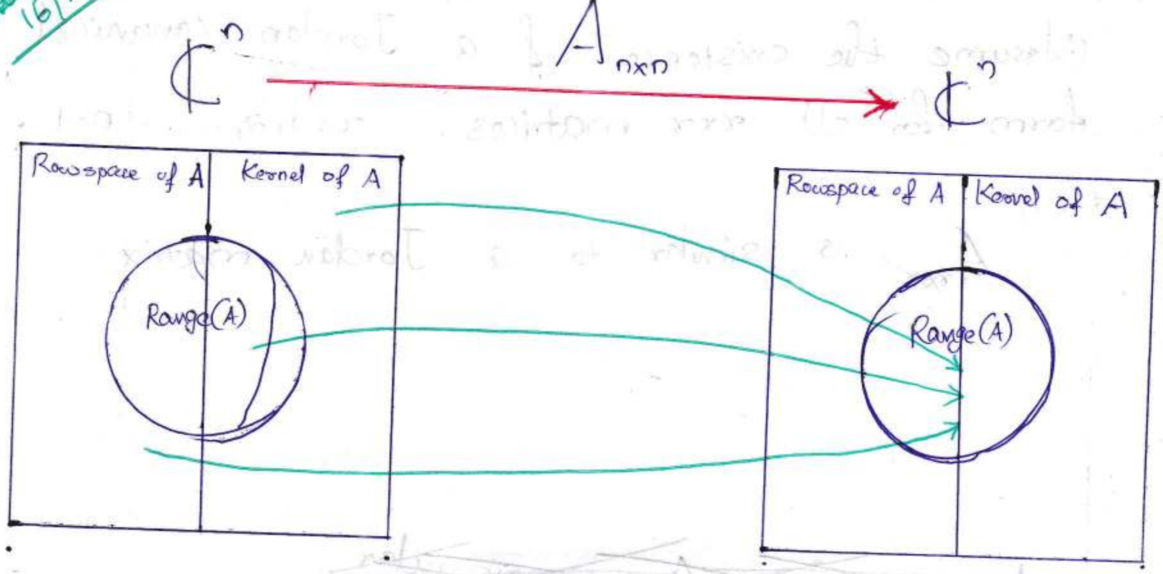
Consider an $n\times n$ matrix $A_{n\times n}$ and assume that $A_{n\times n}$ is singular,

ie.,

$\lambda = 0$ is an eigenvalue

$$\dim\left[\text{range}(A_{n\times n})\right] = r < n$$

~~Check~~ Stack 16/11/2020



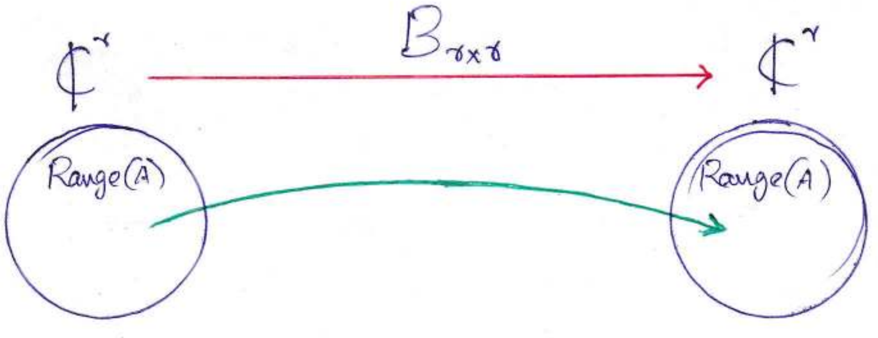
If we think of another transformation, $T : \text{range}(A) \longrightarrow \text{range}(A)$ and the corresponding matrix is $B_{r\times r}$ associated with it.

ie., $B_{r\times r}$ represents the same $\text{range}(A) \rightarrow \text{range}(A)$ transformations as $A_{n\times n}$, just the space becomes smaller,

from the induction hypothesis there exists a Jordan canonical basis $(w_1', \ldots, w_r')$ for the range(A) such that,

$$B w_k' = \lambda_k w_k' \quad (\text{or}) \quad B w_k' = \lambda_k w_k' + w_{k-1}'$$

ie., the linear operator associated with A, restricted by its range has a Jordan canonical form.



Step 1

This implies there exists a Jordan canonical basis $(w_1, \ldots, w_r)$ for the range (A) such that $Aw_k = \lambda_k w_k$ (or) $Aw_k = \lambda_k w_k + w_{k-1}$.

The vectors w_i and w_i' are the same except that w_i is a vector in the larger $\mathbb{C}^n$ space and w_i' is the exact same vector represented in the smaller subspace $\mathbb{C}^r$.

Step 2

Let the subspace $N(A) \cap R(A)$ has dimension p.
The subspace $N(A) \cap R(A)$ is the eigenspace of $R(A)$ corresponding to the eigenvalue $\lambda = 0$. So among the basis vectors $(w_1, \ldots, w_r)$ there are p linearly independent ~~vectors that~~ eigenvectors that has eigenvalue $\lambda = 0$.

Since $w_i \in R(A)$ we have $w_i = A y_i$ for some y_i.

Since $A y_i = 0 y_i + w_i$, we can place each y_i after each corresponding w_i in the string for all p members in $N(A) \cap R(A)$.

~~For scalars a_i, b_i, c_i~~

Assume that there are scalars a_i, b_i, c_i such that

$$\sum_{i=1}^{r} a_i w_i + \sum_{j=1}^{p} b_j y_j + \sum_{k=1}^{n-r-p} c_k z_k = 0 \qquad \text{——— } ①$$

Apply A to both sides, we get

$$\sum_{i} a_i \left(\lambda_i w_i \text{ or } \lambda_i w_i + w_{i-1} \right) + \sum_{j} b_j A y_j = 0 \qquad \text{——— } ②$$

$A z_k = 0$, since $z_k \in N(A)$

$A w_i = 0$ for all $w_i \in N(A) \cap R(A)$

In ②, each Ay_i is one of the w_i in ①.

But,
Ay_i will never coincide with any of the w_i appearing in the 1st summation of ②, because if $w_i = Ay_i$ then $Ay_i = 0$.

By linear independence of the w_i, all the b_j must be zero.

$\Longrightarrow$ $b_j = 0$ for all $j = 1, 2, \cdots, p$

$$\cancel{\sum_{k=1}^{n-r-p} c_k z_k = -\sum_{i=1}^{r} a_i w_i}$$

$$\sum_{i=1}^{r} a_i w_i + \sum_{k=1}^{n-r-p} c_k z_k = 0$$

Here $w_i \in Ran(A)$ and $z_k \notin Ran(A)$, and since they are separately independent, it follows $a_i = 0$, $c_k = 0$. for all i, k.

$\Longrightarrow$ the whole set w_i, y_j, z_k are independent.

$\therefore$ existence of a Jordan Canonical form for every singular square matrix is verified.

$$\cancel{\sum_{k=1}^{n-r-p} c_k z_k = -\sum_{i=1}^{r} a_i w_i}$$

$$\sum_{i=1}^{r} a_i w_i + \sum_{k=1}^{n-r-p} c_k z_k = 0$$

Here $w_i \in Ran(A)$ and $z_k \notin Ran(A)$, and since they are separately independent, it follows $a_i = 0$, $c_k = 0$. for all i, k.

$\Longrightarrow$ the whole set w_i, y_j, z_k are independent.

$\therefore$ existence of a Jordan Canonical form for every singular square matrix is verified.

~~If~~ A is non-singular,

let λ be any eigenvalue of A, and let $A_\lambda = A - \lambda I$, which is singular since $\det(A_\lambda) = 0$.

We have proved that A_λ is similar to a matrix J_0 in Jordan canonical form, so that there is an invertible matrix B such that

$B^{-1} A_\lambda B = J_0$.

$$\cancel{B^{-1} A_\lambda B = B^{-1}(A - \lambda I) B = B^{-1} A B - \lambda I = J_0 - \lambda I}$$

$$B^{-1} A B = B^{-1}(A_\lambda + \lambda I) B = B^{-1} A_\lambda B + \lambda I = J_0 + \lambda I$$

$J = J_0 + \lambda I$ is in Jordan canonical form, (the eigenvalues of A_λ have been increased by λ to produce the eigenvalues of A), and the same basis of generalized eigenvectors that worked for A_λ will work for A, because we are using the same basis changing matrix B.

□ Matrix exponential - nondiagonalizable

A Homogeneous linear systems with constant coefficients

Ж.

$$\frac{d\vec{x}}{dt} = A\vec{x} \implies \vec{x}(t) = e^{At}\vec{x}(0)$$

Any possible solution of $x' = Ax$ can be uniquely expressed in terms of the matrix exponential e^{At}.

e^{At} - identities

$$\left(e^{At}\right)' = A e^{At}$$

$$e^{0} = I$$

Q.C:6

$$e^{At} e^{Bt} = e^{(A+B)t} \quad \text{if } AB = BA$$

$$\left(e^{At}\right)^{-1} = e^{-At}$$

$$e^{At} = \sum_{n=0}^{\infty} \frac{A^n t^n}{n!} = 1 + At + \frac{A^2 t^2}{2!} + \frac{A^3 t^3}{3!} + \cdots\cdots$$

$$= \lim_{n\to\infty}\left(1 + \frac{A}{n}\right)^n$$

~~$A = B^{-1} A B$~~

$$A = BJB^{-1} = B\begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_s \end{bmatrix} B^{-1}$$

$$\longrightarrow \quad e^{At} = B\begin{bmatrix} e^{J_1 t} & 0 & \cdots & 0 \\ 0 & e^{J_2 t} & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & e^{J_s t} \end{bmatrix} B^{-1}$$

where,

$$e^{J_k t} = e^{(\lambda I_k + N_k)t} = e^{\lambda t I_k} \cdot e^{N_k t}$$

$$= e^{\lambda t}\left(I_k + N_k t + \frac{(N_k t)^2}{2} + \frac{(N_k t)^3}{3!} + \cdots + \frac{(N_k t)^{k-1}}{(k-1)!}\right)$$

$$= e^{\lambda t}\begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & & \frac{t^{k-1}}{(k-1)!} \\ 0 & 1 & t & \cdots & & \frac{t^{k-2}}{(k-2)!} \\ 0 & 0 & 1 & \cdots & & \frac{t^{k-3}}{(k-3)!} \\ \vdots & \vdots & & & & \vdots \\ & & & 0 & 1 & t \\ 0 & & & & & \\ 0 & \cdots & & & 0 & 1 \end{bmatrix}$$

where,

N^k : matrix with 1's along the super diagonal.

and taking powers of N^k causes the diagonal of 1 to march up to the right and we get $N_k^k = 0$., thus N_k is nilpotent with k is the degree of nilpotency.

Ex:-

$$J_4 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$= \lambda I_4 + N_4$$

Since λI_4 and N_4 commute,

$$e^{J_4 t} = e^{(\lambda I_4 + N_4)t} = e^{\lambda t I_4} \cdot e^{N_4 t}$$

$$e^{\lambda t I_4} = \begin{bmatrix} e^{\lambda t} & 0 & 0 & 0 \\ 0 & e^{\lambda t} & 0 & 0 \\ 0 & 0 & e^{\lambda t} & 0 \\ 0 & 0 & 0 & e^{\lambda t} \end{bmatrix} = e^{\lambda t} I_4$$

$$e^{N_4 t} = I_4 + N_4 + \frac{(N_4 t)^2}{2!} + \frac{(N_4 t)^3}{3!} + \cdots$$

$$N_4^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad N_4^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$N_4^4 = 0$$

B Non homogeneous linear differential equations

$$x' = Ax + g(t)$$

where, $g(t)$ is a continuous vector valued function.

The general solution of the system can be expressed as:

$$x = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n + x_p(t)$$

where,

$x_1, \ldots, x_n$ is a fundamental set of solutions to the associated homogeneous system $x' = Ax$ and x_p is a particular solution to the non-homogeneous system.

Similar to that for scalar linear equation,

Consider the integrating factor, $\psi(t) = e^{-At}$

$$\frac{dx(t)}{dt} = Ax + g(t) \Longrightarrow e^{-At}\frac{dx}{dt} - e^{-At}Ax = g(t)e^{-At}$$

$$\cancel{\frac{d}{dt}\left(e^{-At}x\right) = g e^{-At}}$$

$$\frac{d}{dt}\left(e^{-At}x\right) = e^{-At}g(t)$$

$$\int_0^t \frac{d}{ds}\left(e^{-As}x\right) ds = \int_0^t e^{-As} g(s)\, ds$$

$$\left[e^{-As}x\right]_0^t = \int_0^t e^{-As} g(s)\, ds$$

$$e^{-At}x(t) - x(0) = \int_0^t e^{-As} g(s)\, ds$$

$$x(t) = e^{At}\int_0^t e^{-As} g(s)\, ds + x(0)\, e^{At}$$

~~Put $t=0$, $x_0 = x(0)$~~

$$\Longrightarrow \quad x(t) = e^{At}x(0) + e^{At}\int_0^t e^{-As} g(s)\, ds$$

$$= x_h(t) + x_p(t)$$

ILA 3

$$\frac{dx}{dt} = Ax + g(t)$$

- Variation of parameters

Ex:- Solve the initial value problem $x' = Ax + g(t)$ with $x(0) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, where $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ & $g(t) = \begin{bmatrix} 2e^{-t} \\ 0 \end{bmatrix}$

Ans:- $A = IJI^{-1}$

$$\exp(At) = \begin{bmatrix} e^{t} & te^{t} \\ 0 & e^{t} \end{bmatrix} = e^{t}\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

$$\exp(-As) = e^{-s}\begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}^{-1} = e^{-s}\begin{bmatrix} 1 & -s \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} e^{-s} & -se^{-s} \\ 0 & e^{-s} \end{bmatrix}$$

$$\int_0^t \exp(-As)\, g(s)\, ds = \int_0^t \begin{bmatrix} e^{-s} & -se^{-s} \\ 0 & e^{-s} \end{bmatrix}\begin{bmatrix} 2e^{-s} \\ 0 \end{bmatrix} ds$$

$$= \int_0^t \begin{bmatrix} 2e^{-2s} \\ 0 \end{bmatrix} ds = \begin{bmatrix} -e^{-2s} \\ 0 \end{bmatrix}_0^t$$

$$= \begin{bmatrix} 1 - e^{-2t} \\ 0 \end{bmatrix}$$

$$x = x_h(t) + x_p(t)$$

$$= x(0)e^{At} + e^{At}\int_0^t e^{-As} g(s)\,ds$$

$$= e^t\begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}\begin{bmatrix}-1\\ 1\end{bmatrix} + e^t\begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}\begin{bmatrix}1-e^{-2t}\\ 0\end{bmatrix}$$

$$= e^t\begin{bmatrix}1 & t\\ 0 & 1\end{bmatrix}\begin{bmatrix}-e^{-2t}\\ 1\end{bmatrix}$$

$$= e^t\begin{bmatrix}t - e^{-2t}\\ 1\end{bmatrix} = \begin{bmatrix}te^t - e^{-t}\\ e^t\end{bmatrix}$$

Q2. Find the eigenvalues and all possible Jordan forms if $A^2 = 0$.

Sol:
$$Ax = \lambda x \longrightarrow A^2 x = \lambda^2 x = 0x$$

$\therefore \underline{\lambda = 0}$.

$$J^2 = (B^{-1}AB)(B^{-1}AB) = B^{-1}A^2B = 0.$$

Every block in J has $\lambda = 0$ on the diagonal.

J_k^2 for block sizes 1, 2, 3:

$$[0]^2 = [0], \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$J^2 = 0 \Longrightarrow$ all block sizes must be 1 or 2.
$J^2 \neq 0$ for 3×3.

④ $B_{in} = B_{out} =$ Fourier matrix, F

$\Longrightarrow$ Fx is a Discrete Fourier Transform of x.

Which matrices are diagonalized by F ?

Starting with the eigenvectors $(1, \lambda, \lambda^2, \lambda^3)$ and finding the matrices that have those eigenvectors:

$\lambda^4 = 1$

$$Px = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \\ \lambda^3 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ \lambda \\ \lambda^2 \\ \lambda^3 \end{bmatrix} = \lambda x$$

P : permutation matrix

4th row of this vector equation is $1 = \lambda^4$.

$\lambda^4 = 1 \Longrightarrow \lambda = 1, i, -1, -i$

all these are eigenvalues of P, each with its eigenvector $x = (1, \lambda, \lambda^2, \lambda^3)$.

The eigenvector matrix F diagonalizes the permutation matrix P:

Eigenvalue matrix Λ : $\begin{bmatrix} 1 & & & \\ & i & & \\ & & -1 & \\ & & & -i \end{bmatrix}$

Eigenvector matrix is Fourier matrix F : $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & i^2 & 1 & (-i)^2 \\ 1 & i^3 & -1 & (-i)^3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & +i \end{bmatrix}$

The columns of F are orthogonal (unitary) because they are eigenvectors of P (an orthogonal (unitary) matrix)

* F is the most important complex matrix in the world.

What other matrices beyond P have this same eigenvector matrix F ?

P^2, P^3, P^4 have the same eigenvectors as P.
The same matrix F diagonalizes all powers of P.

If $P, P^2, P^3, P^4 = I$ have the same eigenvector matrix F, so does any combination

$$C = c_1 P + c_2 P^2 + c_3 P^3 + c_0 P^4$$
$$= c_1 P + c_2 P^2 + c_3 P^4 + c_0 I$$

Circulant matrix,

$$C = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_3 & c_0 & c_1 & c_2 \\ c_2 & c_3 & c_0 & c_1 \\ c_1 & c_2 & c_3 & c_0 \end{bmatrix}$$

has eigenvectors in the Fourier matrix F.
has 4 eigenvalues $c_0 + c_1 \lambda + c_2 \lambda^2 + c_3 \lambda^3$, from the 4 numbers $\lambda = 1, i, -1, -i$.

The 4 eigenvalues of C are given by the Fourier transform Fc,

$$Fc = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} c_0 + c_1 + c_2 + c_3 \\ c_0 + ic_1 - c_2 - ic_3 \\ c_0 - c_1 + c_2 - c_3 \\ c_0 - ic_1 - c_2 + ic_3 \end{bmatrix}$$

- Circulant matrices have constant diagonals. The same number c_0 goes down the main diagonal. The number c_1 is on the diagonal above, and that diagonal "wraps around" or "circles around" to the southwest corner of C.

Bases for Function Space

For functions of x, the 1st basis I'd think of contains the powers $1, x, x^2, x^3, \ldots$.
Unfortunately this is a terrible basis.
Those functions x^n are just barely independent
x^{10} is almost a combination of other basis vectors $1, x, \ldots, x^9$. It is virtually impossible to compute with this poor "ill-conditioned" basis.

If we had vectors instead of functions, the test for a good basis would look at $B^T B$. This matrix contains all inner products b/w the basis vectors (columns of B). The basis is orthonormal when $B^T B = I$. That is best possible
But the basis $1, x, x^2, \ldots$ produces the evil Hilbert matrix: $B^T B$ has an enormous ratio b/w its largest and smallest eigenvalues.

$\longrightarrow$ A large condition number signals an unhappy choice of basis.

Why is $1, x, x^2, x^3, \ldots$ a bad basis?

Stack
6/12/2020 $\longrightarrow$

The vectors $(1,0,0), (1,0.01,0), (1,0.01,0.01)$ are linearly independent in $\mathbb{R}^3$, but are very close to each other.

If you are working in the space of functions defined over $[-1,1]$

Strang says $\Rightarrow$ projection $p(x)$ of x^{10} onto the space spanned by $\{1, x, x^2, \ldots, x^9\}$ is quite close to x^{10} itself in a least square sense.

$$\|p(x) - x^{10}\| = \int_{-1}^{1} |p(x) - x^{10}|^2 dx \quad \text{is small}$$

$$\|x^8 - x^{10}\| = \int_{-1}^{1} x^{16}(1-x^2)^2 dx = \int_{-1}^{1} (x^{16} - 2x^{18} + x^{20}) dx$$

$$\approx 0.00235$$

$$\|1 - x^{10}\|^2 = \int_{-1}^{+1} (1-x^{10})^2 dx = \int_{-1}^{1} (1 - 2x^{10} + x^{20}) dx = \left[x - \frac{2x^{11}}{11} + \frac{x^{21}}{21}\right]_{-1}^{1}$$

$$\approx 1.9134$$

The square distance b/w $p(x)$ and x^{10} is upper bounded by this value. And, the distance b/w x^k and the least squares approximation of x^k restricted to the subspace spanned by $\{1, x, \ldots, x^{k-1}\}$ will tend to zero as $k \to \infty$. ie., intuitively, this means that the higher and higher k goes, the closer and closer you get to adding a linearly dependent vector to your set.

Now the columns of B are functions instead of vectors. We still use $B^T B$ to test for independence. So we need to know the dot product (inner product) of 2 functions — those are the numbers in $B^T B$.

$$x^T y = x_1 y_1 + \cdots\cdots + x_n y_n$$

$$\text{Inner product } (f, g) = \int f(x) g(x)\, dx$$

$$\text{Complex inner product } (f, g) = \int \overline{f(x)}\, g(x)\, dx,$$

$\bar{f}$: complex conjugate

$$\text{Weighted inner product } (f, g)_w = \int w(x)\, \overline{f(x)}\, g(x)\, dx,$$

w : weight function.

When the integrals go from $x=0$ to $x=1$, the inner product of x^i and x^j is:

$$\int_0^1 x^i x^j \, dx = \left. \frac{x^{i+j+1}}{i+j+1} \right]_0^1 = \frac{1}{i+j+1} = \text{entries of Hilbert matrix } B^T B.$$

By changing to the symmetric interval from $x=-1$ to $x=1$, we have orthogonality b/w all even functions and all odd functions:

Interval $[-1, 1]$: $\displaystyle\int_{-1}^{1} x^2 x^5 \, dx = 0$

$$\int_{-1}^{1} \text{even}(x) \, \text{odd}(x) \, dx = 0$$

This change makes half of the basis functions orthogonal to the other half. It is so simple that we continue using the symmetric interval -1 to 1 (or $-\pi$ to π). But we want a better basis than the powers x^n — hopefully an orthogonal basis.

Hilbert Space

After studying $\mathbb{R}^n$, it is natural to think of the space $\mathbb{R}^\infty$. It contains all vectors $v = (v_1, v_2, \dots)$ with an infinite sequence of components.
This space is actually too big when there is no control on the size of components v_j.

A much better idea is to keep the familiar definition of length, using a sum of squares, and to include only those vectors that have a finite length:

Length squared: $\|v\|^2 = v_1^2 + v_2^2 + v_3^2 + \cdots$

The infinite series must converge to a finite sum.

~~The~~
Vectors with finite length can be added $(\|v+w\| \le \|v\| + \|w\|)$ and multiplied by scalars, so they form a vector space. It is the Hilbert space.

Hilbert space is the natural way to let the number of dimensions become infinite, and at the same time to keep the geometry of ordinary Euclidean space.

Hilbert Space

After studying $\mathbb{R}^n$, it is natural to think of the space $\mathbb{R}^\infty$. It contains all vectors $v = (v_1, v_2, \dots)$ with an infinite sequence of components.
This space is actually too big when there is no control on the size of components v_j.

A much better idea is to keep the familiar definition of length, using a sum of squares, and to include only those vectors that have a finite length:

Length squared: $\|v\|^2 = v_1^2 + v_2^2 + v_3^2 + \dots$

The infinite series must converge to a finite sum.

~~The~~
Vectors with finite length can be added $(\|v+w\| \leq \|v\| + \|w\|)$ and multiplied by scalars, so they form a vector space. It is the Hilbert space.

Hilbert space is the natural way to let the number of dimensions become infinite, and at the same time to keep the geometry of ordinary Euclidean space.

Orthogonality : $v^T w = v_1 w_1 + v_2 w_2 + \cdots\cdots \cancel{} = 0$

This sum is guaranteed to converge, and for any 2 vectors it still obeys the Schwarz inequality $|v^T w| \leq \|v\| \cdot \|w\|$

There is another remarkable thing about this space: It is found under a great many different disguises. Its "vectors" can turn into functions.

Say, $f(x) = \sin(x)$ on the interval $0 \le x \le 2\pi$.

This "f" is like a vector with a whole continuum of components, the values of $\sin(x)$ along the whole interval. To find the length of such a vector, the usual rule of adding the squares of the components become impossible. This summation is replaced in a natural and inevitable way by integration:

Length $\|f\|$ of function : $\|f\|^2 = \int_0^{2\pi} (f(x))^2 dx = \int_0^{2\pi} (\sin x)^2 dx = \pi$

$\Longrightarrow$ Our Hilbert space has become a functional space. The vectors are functions, we have a way to measure their lengths, and the space contains all those functions that have a finite length. It does not contain the function $F(x) = \frac{1}{x}$, because the integral of $\frac{1}{x^2}$ is ∞.

If $f(x) = \sin x$ and $g(x) = \cos x$, their inner product is :

$$(f, g) = \int_0^{2\pi} f(x)\, g(x)\, dx = \int_0^{2\pi} \sin x \cos x = 0$$

This is exactly like the vector inner product $f^T g$. It is still related to the length by $(f, f) = \|f\|^2$. The Schwarz inequality is still satisfied :

$|(f, g)| \leq \|f\| \|g\|$.

→ 2 functions like $\sin x$, $\cos x$ whose inner product is zero, will be called orthogonal. They are even orthonormal after division by their length $\sqrt{\pi}$.

* The vector $v = (v_1, v_2, \cdots\cdots)$ and the function $f(x)$ are in our infinite dimensional "Hilbert spaces" iff their lengths $\|v\|$ and $\|f\|$ are finite:

$$\|v\|^2 = v \cdot v = v_1^2 + v_2^2 + v_3^2 + \cdots\cdots$$ must add to a finite #.

$$\|f\|^2 = (f \cdot f) = \int_0^{2\pi} |f(x)|^2 \, dx$$ must be a finite integral.

$v = \left(1, \frac{1}{2}, \frac{1}{4}, \cdots\right)$ is included in Hilbert space

because,

$$v \cdot v = 1 + \frac{1}{4} + \frac{1}{16} + \cdots\cdots = \frac{1}{1 - \frac{1}{4}} = \frac{4}{3}$$