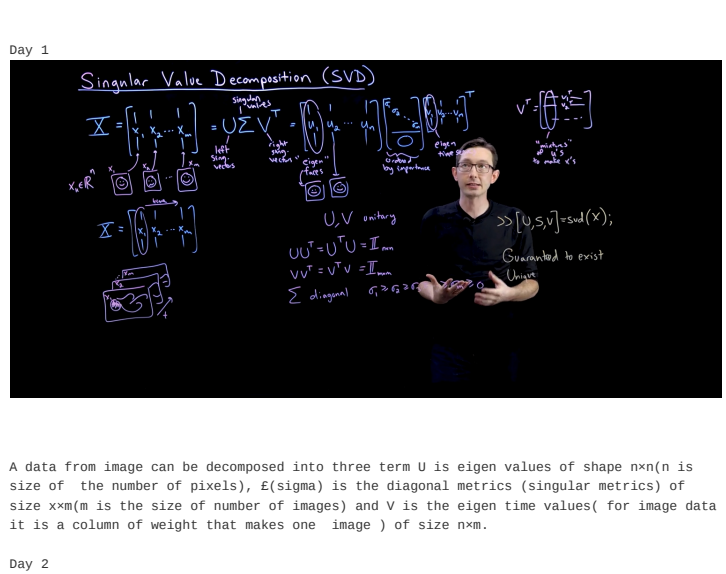
Last updated: Nov 12, 2020

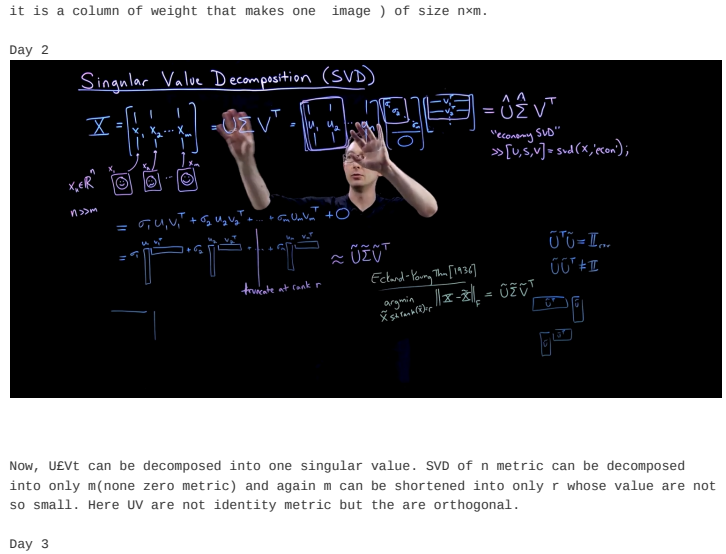
SVD STEVEN BURNTON

Day 1


Singular Value Decomposition (SVD)
U,V unitary
Guaranteed to exist
Unique
>>[U,S,V]=svd(X);

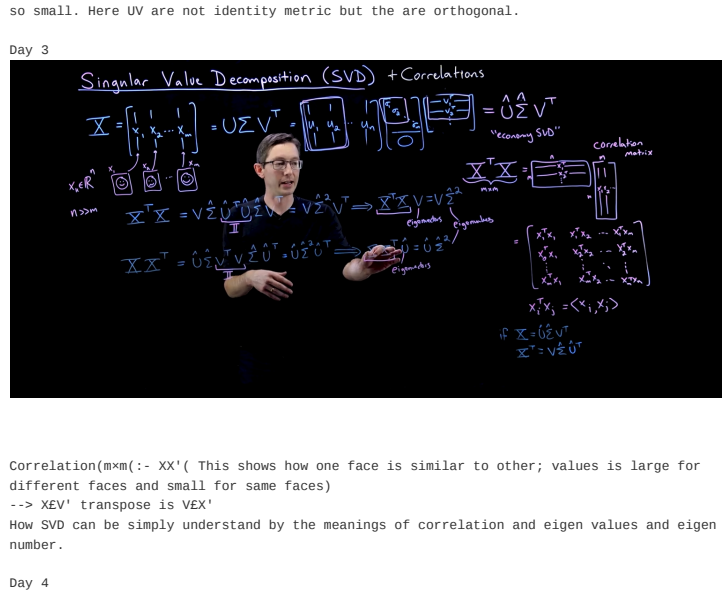
A data from image can be decomposed into three term U is eigen values of shape n×n(n is size of the number of pixels), £(sigma) is the diagonal metrics (singular metrics) of size x×m(m is the size of number of images) and V is the eigen time values(for image data it is a column of weight that makes one image) of size n×m.

Day 2


Singular Value Decomposition (SVD)
"economy SVD"
>>[U,S,V]=svd(X,'econ');
Eckard-Young Thm [1936]
truncate at rank r

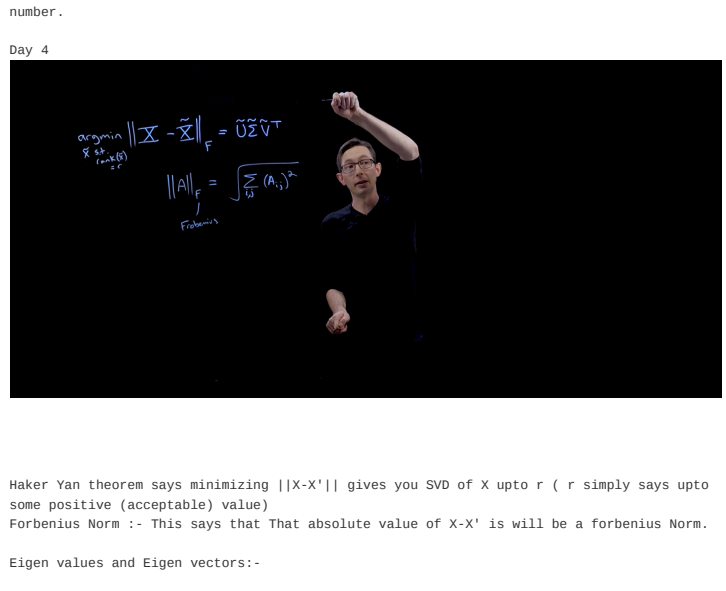
Now, U£Vt can be decomposed into one singular value. SVD of n metric can be decomposed into only m(none zero metric) and again m can be shortened into only r whose value are not so small. Here UV are not identity metric but the are orthogonal.

Day 3


Singular Value Decomposition (SVD) + Correlations
"economy SVD"
correlation matrix
eigenvectors
eigenvalues

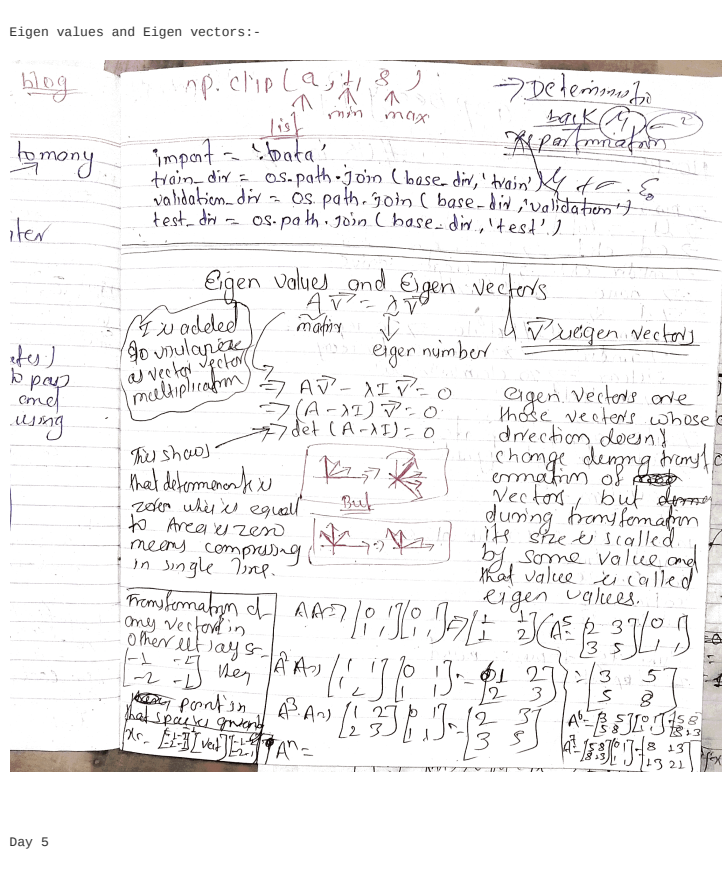
Correlation(m×m(:- XX'(This shows how one face is similar to other; values is large for different faces and small for same faces)
--> X£V' transpose is V£X'
How SVD can be simply understand by the meanings of correlation and eigen values and eigen number.

Day 4


argmin ||X - X̃||_F = ŨΣ̃Ṽ^T
Frobenius

Haker Yan theorem says minimizing ||X-X'|| gives you SVD of X upto r (r simply says upto some positive (acceptable) value)
Forbenius Norm :- This says that That absolute value of X-X' is will be a forbenius Norm.

Eigen values and Eigen vectors:-



np.clip(a, 1, 8)

list

min max

→ Determinant

back

Transformation

import – 'Data'

train_dir = os.path.join(base_dir,'train')

validation_dir = os.path.join(base_dir,'validation')

test_dir = os.path.join(base_dir,'test')

Eigen values and Eigen vectors

$A\vec{v} = \lambda\vec{v}$

matrix

eigen number

$\vec{v}$ = eigen vectors

I is added to vularize a vector vector multiplication

$\Rightarrow A\vec{v} - \lambda I\vec{v} = 0$

$\Rightarrow (A-\lambda I)\vec{v} = 0$

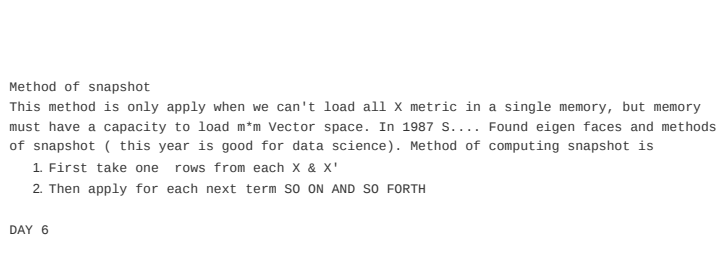
$\Rightarrow \det(A-\lambda I) = 0$

This shows that determinant is zero which is equal to Area is zero means compressing in single line.

eigen vectors are those vectors whose direction doesn't change during transformation of vector, but during transformation its size is scalled by some value and that value is called eigen values.

Transformation of any vector in other ways

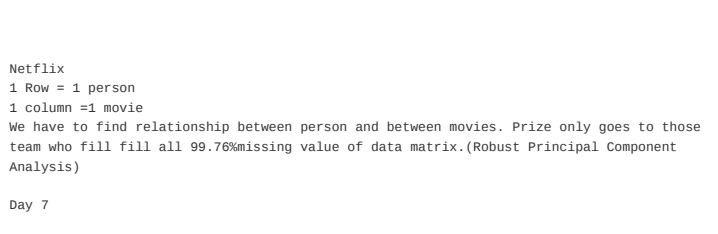
Day 5


Singular Value Decomposition (SVD) + Correlations
"economy SVD"
Method of snapshots

Method of snapshot
This method is only apply when we can't load all X metric in a single memory, but memory must have a capacity to load m*m Vector space. In 1987 S.... Found eigen faces and methods of snapshot (this year is good for data science). Method of computing snapshot is

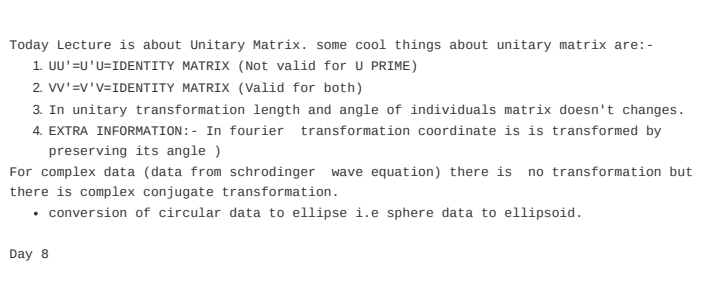
1. First take one rows from each X & X'
2. Then apply for each next term SO ON AND SO FORTH

DAY 6


Singular Value Decomposition (SVD)
Netflix Prize
RPCA
(ch3)

Netflix
1 Row = 1 person
1 column =1 movie
We have to find relationship between person and between movies. Prize only goes to those team who fill fill all 99.76%missing value of data matrix.(Robust Principal Component Analysis)

Day 7


Unitary Transformations & Geometry
Preserve angles and lengths of vectors
Just rotate vectors
If X complex

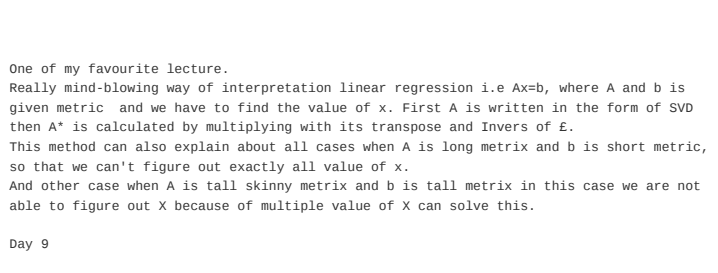
Today Lecture is about Unitary Matrix. some cool things about unitary matrix are:-

1. UU'=U'U=IDENTITY MATRIX (Not valid for U PRIME)
2. VV'=V'V=IDENTITY MATRIX (Valid for both)
3. In unitary transformation length and angle of individuals matrix doesn't changes.
4. EXTRA INFORMATION:- In fourier transformation coordinate is is transformed by preserving its angle)

For complex data (data from schrodinger wave equation) there is no transformation but there is complex conjugate transformation.

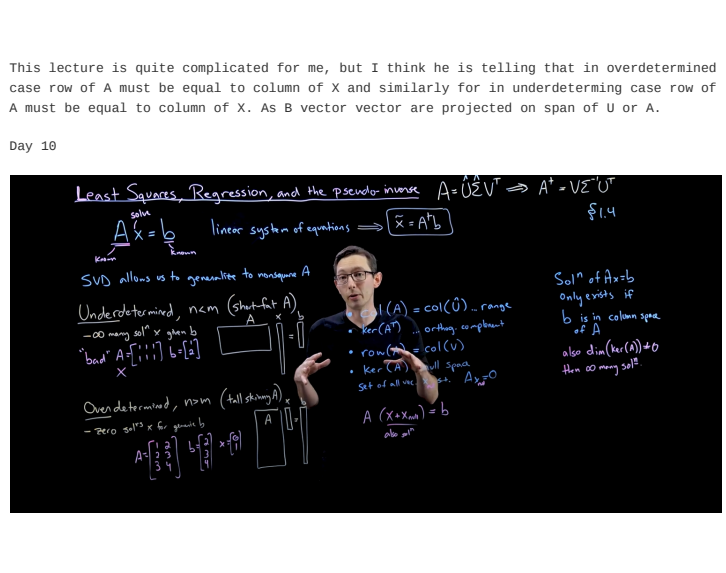
- conversion of circular data to ellipse i.e sphere data to ellipsoid.

Day 8


Least Squares, Regression, and the pseudo-inverse
linear system of equations
Underdetermined
Overdetermined

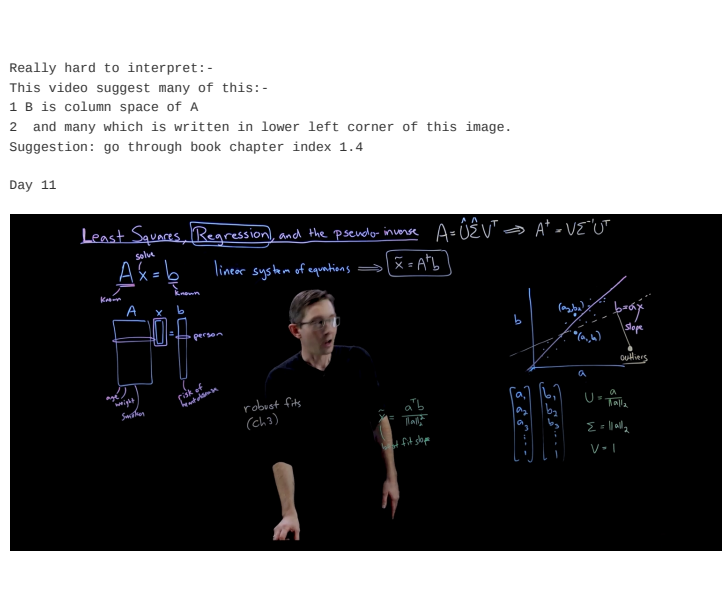
One of my favourite lecture.
Really mind-blowing way of interpretation linear regression i.e Ax=b, where A and b is given metric and we have to find the value of x. First A is written in the form of SVD then A+ is calculated by multiplying with its transpose and Invers of £.
This method can also explain about all cases when A is long metrix and b is short metric, so that we can't figure out exactly all value of x.
And other case when A is tall skinny metrix and b is tall metrix in this case we are not able to figure out X because of multiple value of X can solve this.

Day 9


Least Squares, Regression, and the pseudo-inverse

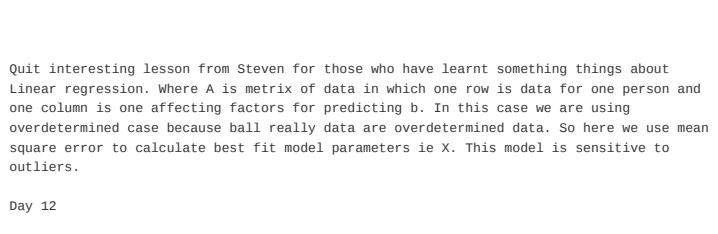
This lecture is quite complicated for me, but I think he is telling that in overdetermined case row of A must be equal to column of X and similarly for in underdeterming case row of A must be equal to column of X. As B vector vector are projected on span of U or A.

Day 10


Least Squares, Regression, and the pseudo-inverse

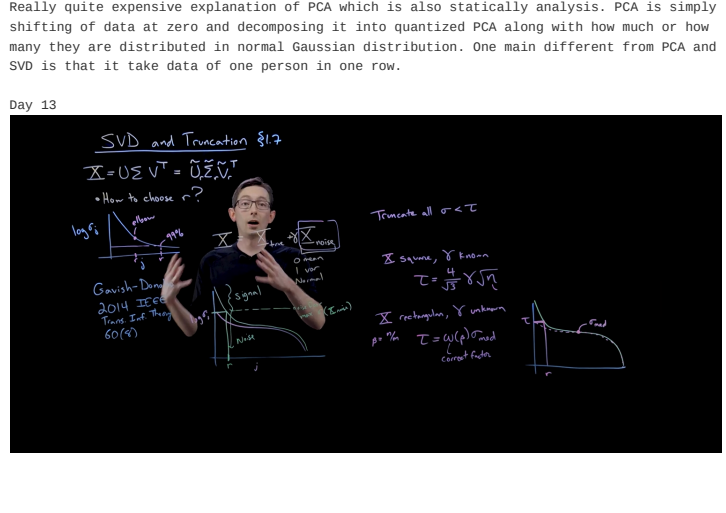
Really hard to interpret:-
This video suggest many of this:-
1 B is column space of A
2 and many which is written in lower left corner of this image.
Suggestion: go through book chapter index 1.4

Day 11


Least Squares, Regression, and the pseudo-inverse

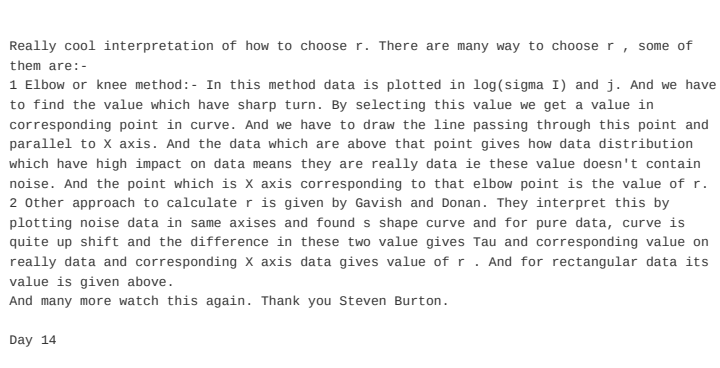
Quit interesting lesson from Steven for those who have learnt something things about Linear regression. Where A is metrix of data in which one row is data for one person and one column is one affecting factors for predicting b. In this case we are using overdetermined case because ball really data are overdetermined data. So here we use mean square error to calculate best fit model parameters ie X. This model is sensitive to outliers.

Day 12


Principal Component Analysis (PCA)

Really quite expensive explanation of PCA which is also statically analysis. PCA is simply shifting of data at zero and decomposing it into quantized PCA along with how much or how many they are distributed in normal Gaussian distribution. One main different from PCA and SVD is that it take data of one person in one row.

Day 13


SVD and Truncation
How to choose r?
Gavish-Donoho 2014

Really cool interpretation of how to choose r. There are many way to choose r , some of them are:-
1 Elbow or knee method:- In this method data is plotted in log(sigma I) and j. And we have to find the value which have sharp turn. By selecting this value we get a value in corresponding point in curve. And we have to draw the line passing through this point and parallel to X axis. And the data which are above that point gives how data distribution which have high impact on data means they are really data ie these value doesn't contain noise. And the point which is X axis corresponding to that elbow point is the value of r.
2 Other approach to calculate r is given by Gavish and Donan. They interpret this by plotting noise data in same axises and found s shape curve and for pure data, curve is quite up shift and the difference in these two value gives Tau and corresponding value on really data and corresponding X axis data gives value of r . And for rectangular data its value is given above.
And many more watch this again. Thank you Steven Burton.

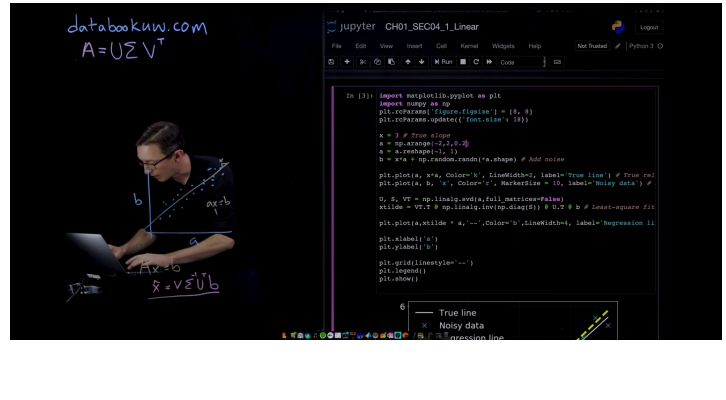
Day 14

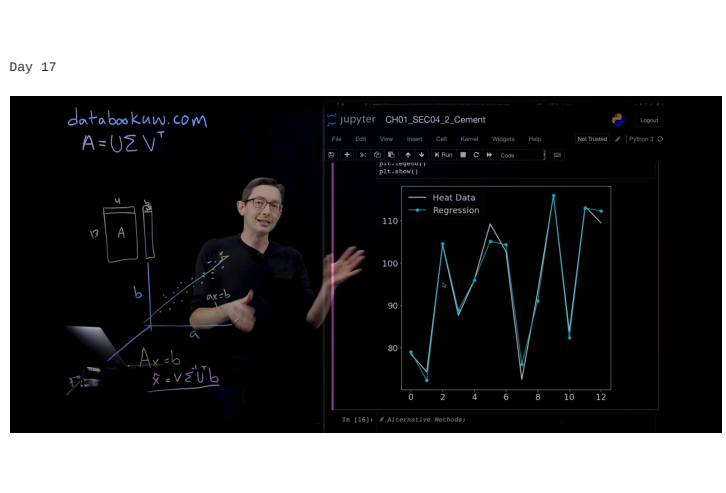
Really nice way of reducing data matrix using SVD, MUST TRY ipynb or .m file.

Day 15

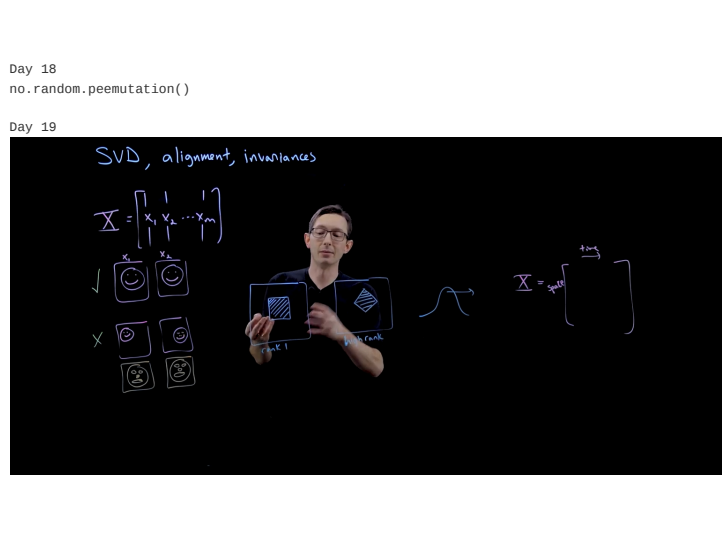
MATLAB transfer of identity metric. Ie conversion of sphere to ellipsoid.

Day 16
Linear Regression
SVD


databookuw.com
jupyter CH01_SEC04_1_Linear

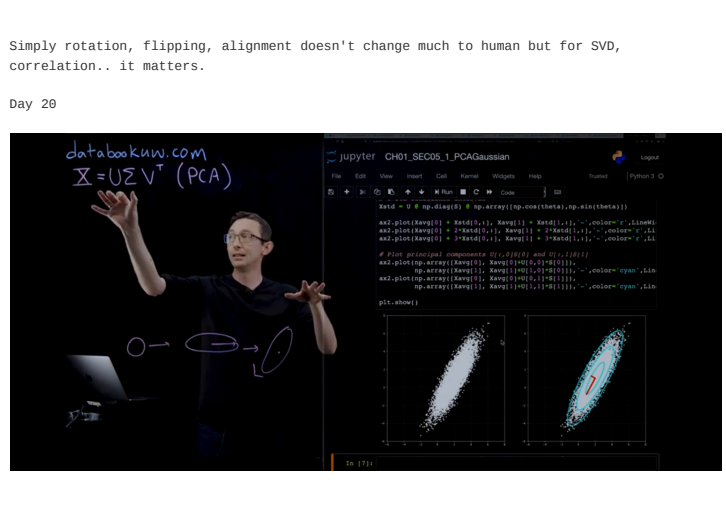

databookuw.com
jupyter CH01_SEC04_1_Linear
True line
Noisy data

Day 17


databookuw.com
jupyter CH01_SEC04_2_Cement
Heat Data
Regression

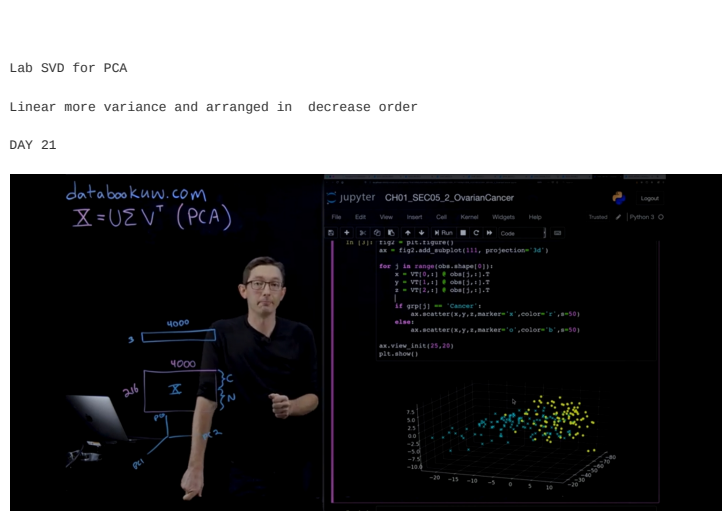
Day 18
no.random.peemutation()

Day 19


SVD, alignment, invariances

Simply rotation, flipping, alignment doesn't change much to human but for SVD, correlation.. it matters.

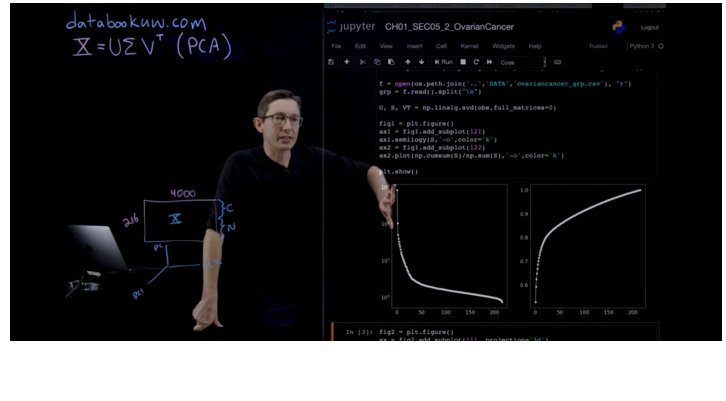
Day 20

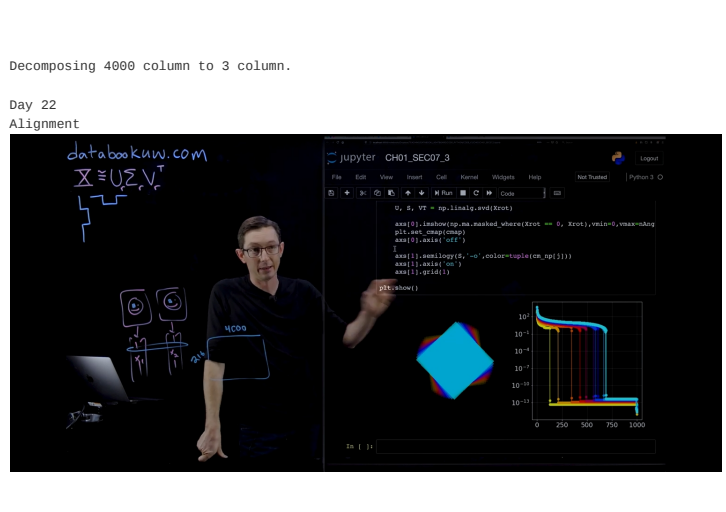

databookuw.com
jupyter CH01_SEC05_1_PCAGaussian

Lab SVD for PCA

Linear more variance and arranged in decrease order

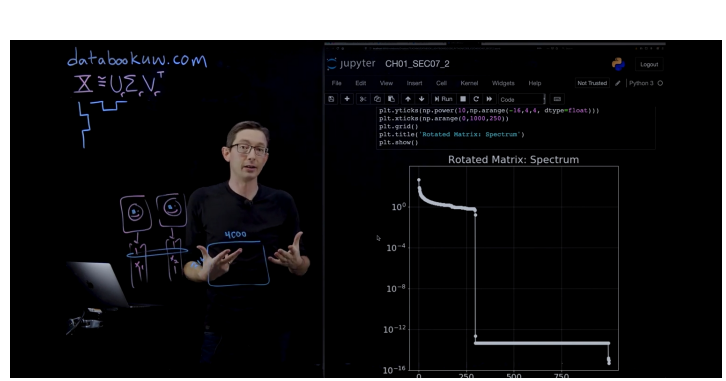
DAY 21

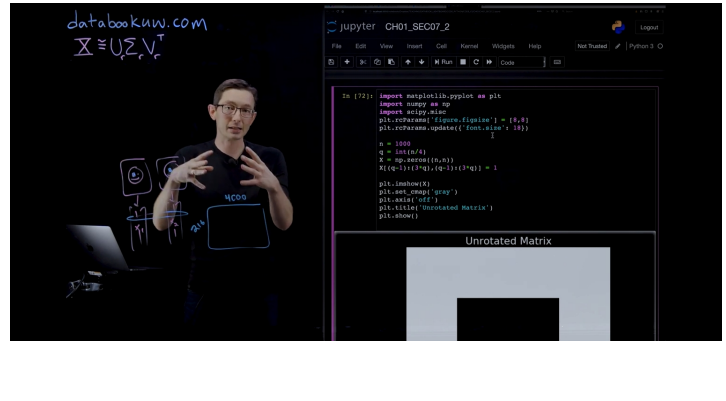

databookuw.com
jupyter CH01_SEC05_2_OvarianCancer

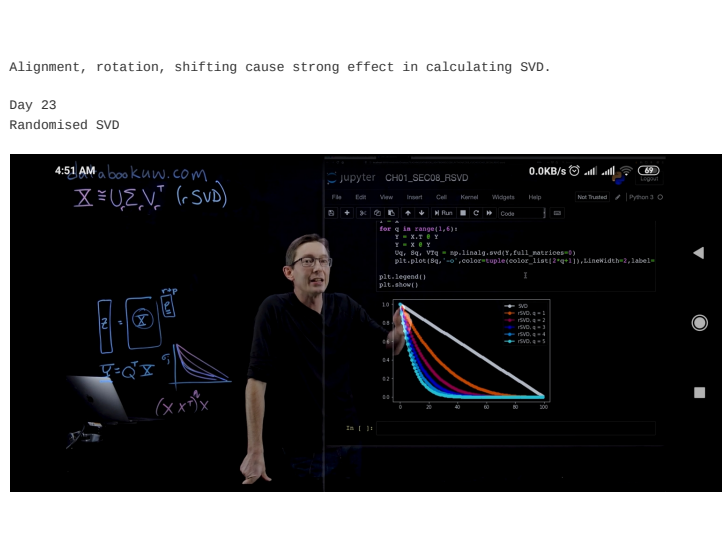

databookuw.com
jupyter CH01_SEC05_2_OvarianCancer

Decomposing 4000 column to 3 column.

Day 22
Alignment

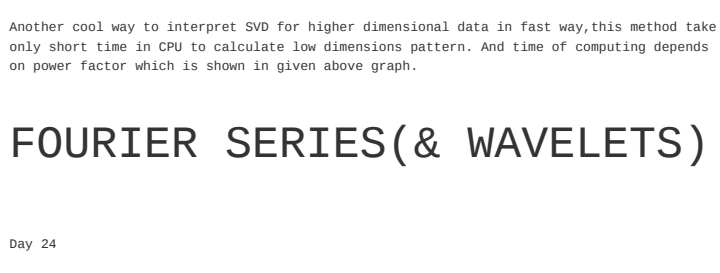

databookuw.com
jupyter CH01_SEC07_1


databookuw.com
jupyter CH01_SEC07_2
Rotated Matrix: Spectrum


databookuw.com
jupyter CH01_SEC07_2
Unrotated Matrix

Alignment, rotation, shifting cause strong effect in calculating SVD.

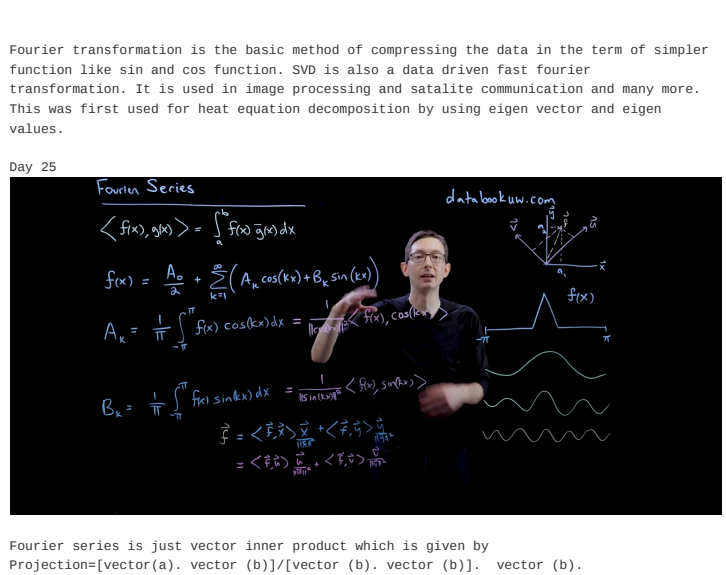
Day 23
Randomised SVD


databookuw.com
jupyter CH01_SEC08_RSVD

Another cool way to interpret SVD for higher dimensional data in fast way,this method take only short time in CPU to calculate low dimensions pattern. And time of computing depends on power factor which is shown in given above graph.

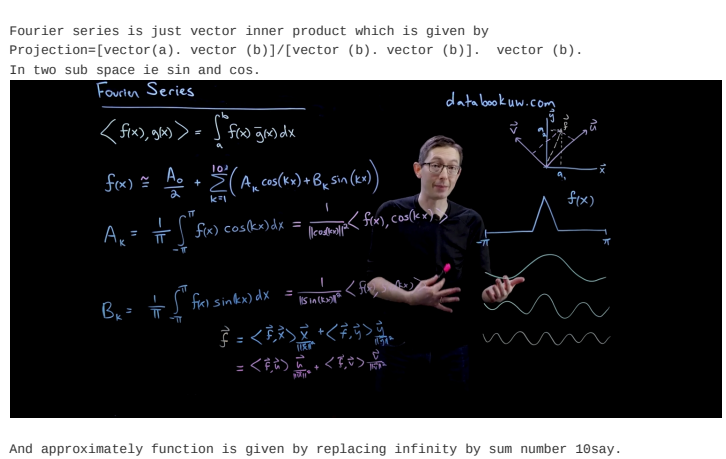
FOURIER SERIES(& WAVELETS)

Day 24

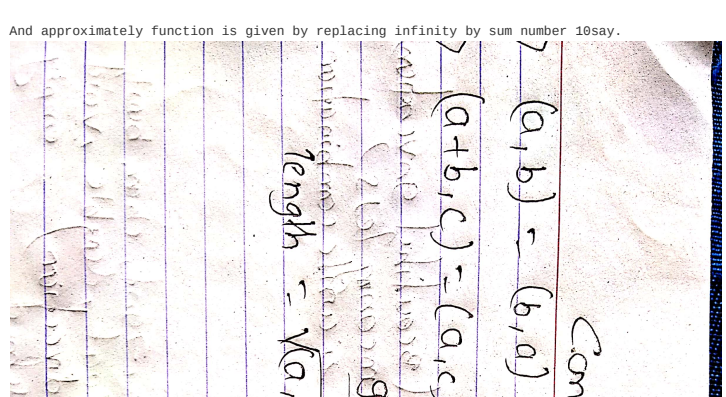

Fourier Transform (& wavelets)
databookuw.com
Coordinate transform
SVD = Data-driven FFT
Hilbert
Fast Fourier Transf. (FFT)

Fourier transformation is the basic method of compressing the data in the term of simpler function like sin and cos function. SVD is also a data driven fast fourier transformation. It is used in image processing and satalite communication and many more. This was first used for heat equation decomposition by using eigen vector and eigen values.

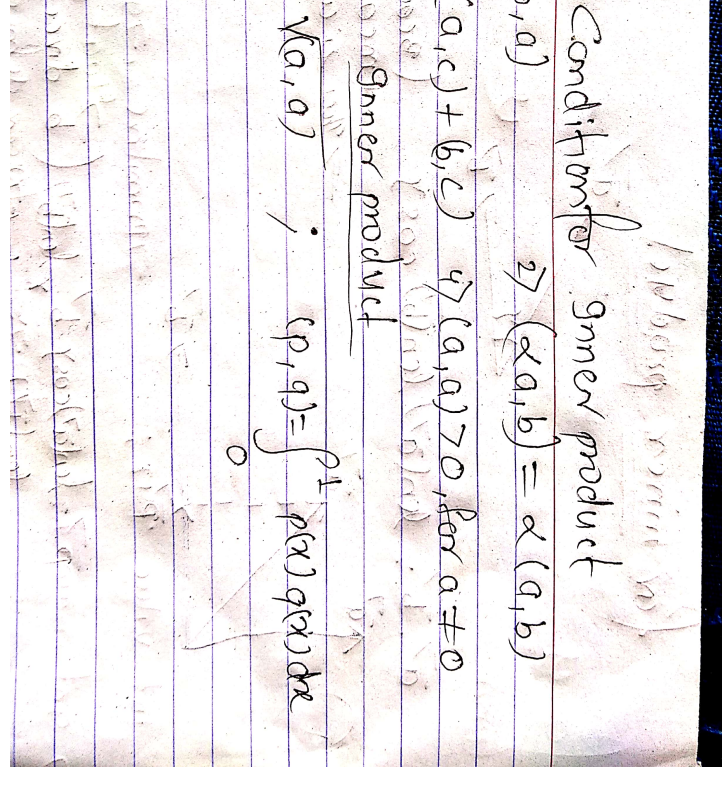
Day 25


Fourier Series
databookuw.com

Fourier series is just vector inner product which is given by
Projection=[vector(a). vector (b)]/[vector (b). vector (b)]. vector (b).
In two sub space ie sin and cos.


Fourier Series
databookuw.com

And approximately function is given by replacing infinity by sum number 10say.



Condition for inner product

1) $\langle a,b\rangle = \langle b,a\rangle$

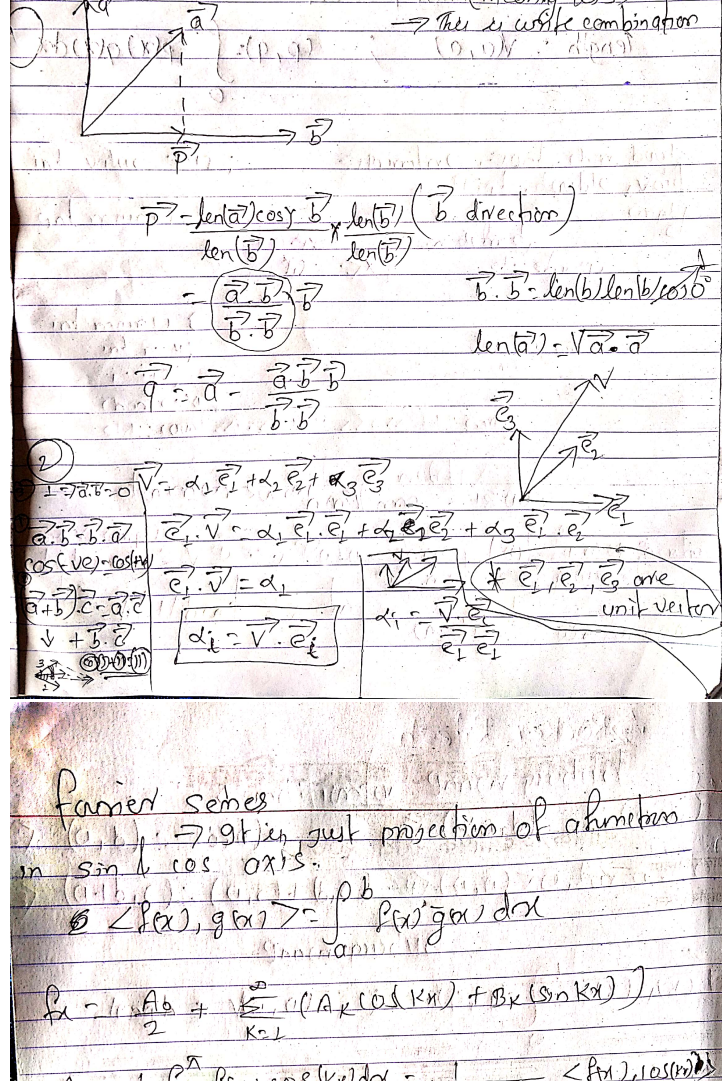
2) $\langle \alpha a,b\rangle = \alpha\langle a,b\rangle$

3) $\langle a+b,c\rangle = \langle a,c\rangle + \langle b,c\rangle$

4) $\langle a,a\rangle > 0$, for $a \neq 0$

Inner product

length $= \sqrt{\langle a,a\rangle}$; $\langle p,q\rangle = \int_0^1 p(x)q(x)\,dx$



for inner product

$\vec{a}\cdot\vec{b} = \mathrm{len}(\vec{a})\,\mathrm{len}(\vec{b})\cos\gamma$ (Beautiful convention) (meaningless)

→ This is write combination

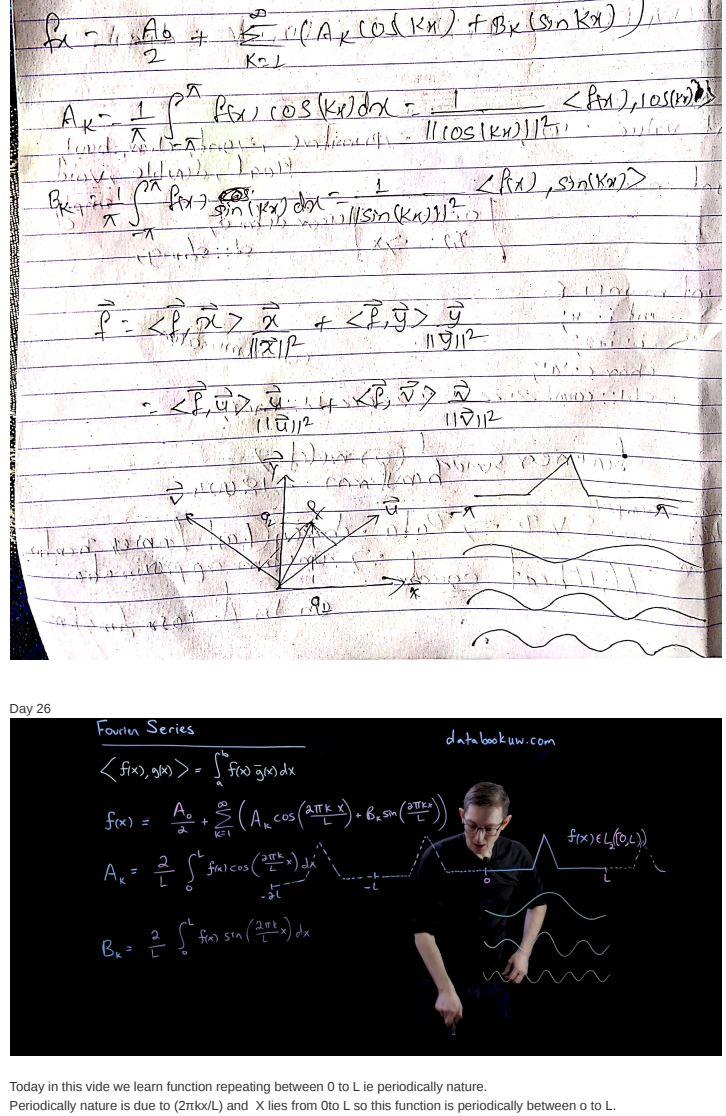
$\vec{p} = \mathrm{len}(\vec{a})\cos\gamma \, \frac{\vec{b}}{\mathrm{len}(\vec{b})}$ ($\vec{b}$ direction)

$\vec{b}\cdot\vec{b} = \mathrm{len}(b)\,\mathrm{len}(b)\cos 0°$

$\mathrm{len}(\vec{a}) = \sqrt{\vec{a}\cdot\vec{a}}$

$\vec{e}_1, \vec{e}_2, \vec{e}_3$ are unit vector

$\alpha_1 = \vec{V}\cdot\vec{e}_1$



Fourier series

Fourier series is just projection of a function in sin & cos axis.

$\langle f(x), g(x)\rangle = \int_a^b f(x)\,g(x)\,dx$

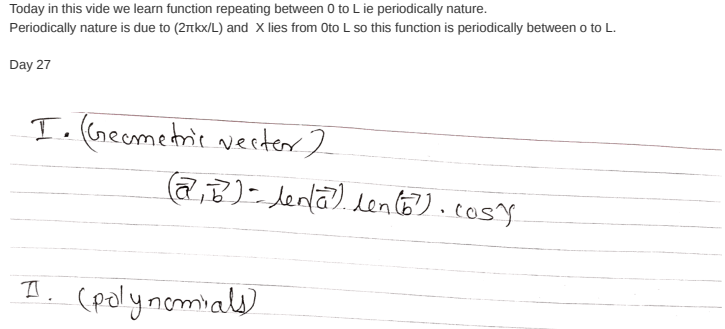
$f(x) = \frac{A_0}{2} + \sum_{k=1}^{\infty}\left(A_k\cos(kx) + B_k\sin(kx)\right)$

$A_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)\,dx = \frac{1}{\|\cos(kx)\|^2}\langle f(x), \cos(kx)\rangle$

$B_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)\,dx = \frac{1}{\|\sin(kx)\|^2}\langle f(x), \sin(kx)\rangle$

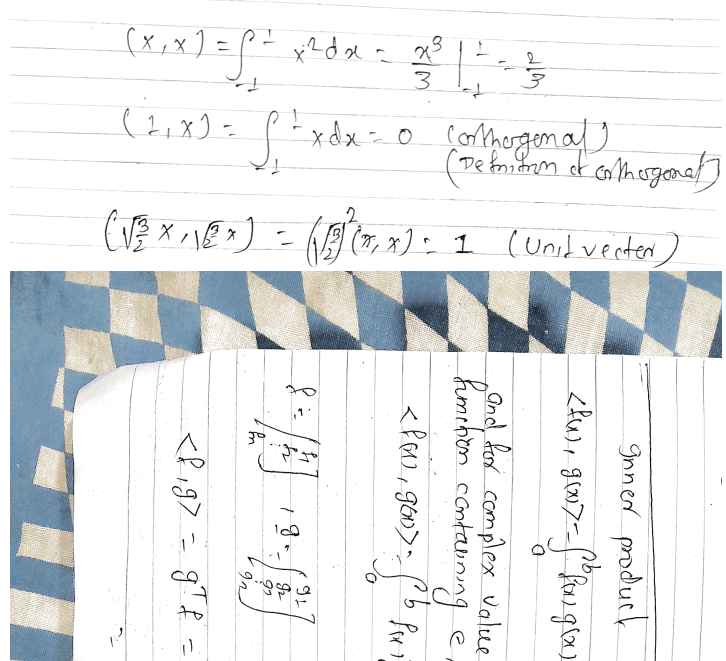
$\vec{f} = \frac{\langle \vec{f},\vec{x}\rangle}{\|\vec{x}\|^2}\vec{x} + \frac{\langle \vec{f},\vec{y}\rangle}{\|\vec{y}\|^2}\vec{y}$

Day 26


Fourier Series
databookuw.com

Today in this vide we learn function repeating between 0 to L ie periodically nature.
Periodically nature is due to (2πkx/L) and X lies from 0to L so this function is periodically between o to L.

Day 27



I. (Geometric vector)

$\langle\vec{a},\vec{b}\rangle = \mathrm{len}(\vec{a})\,\mathrm{len}(\vec{b})\cdot\cos\gamma$

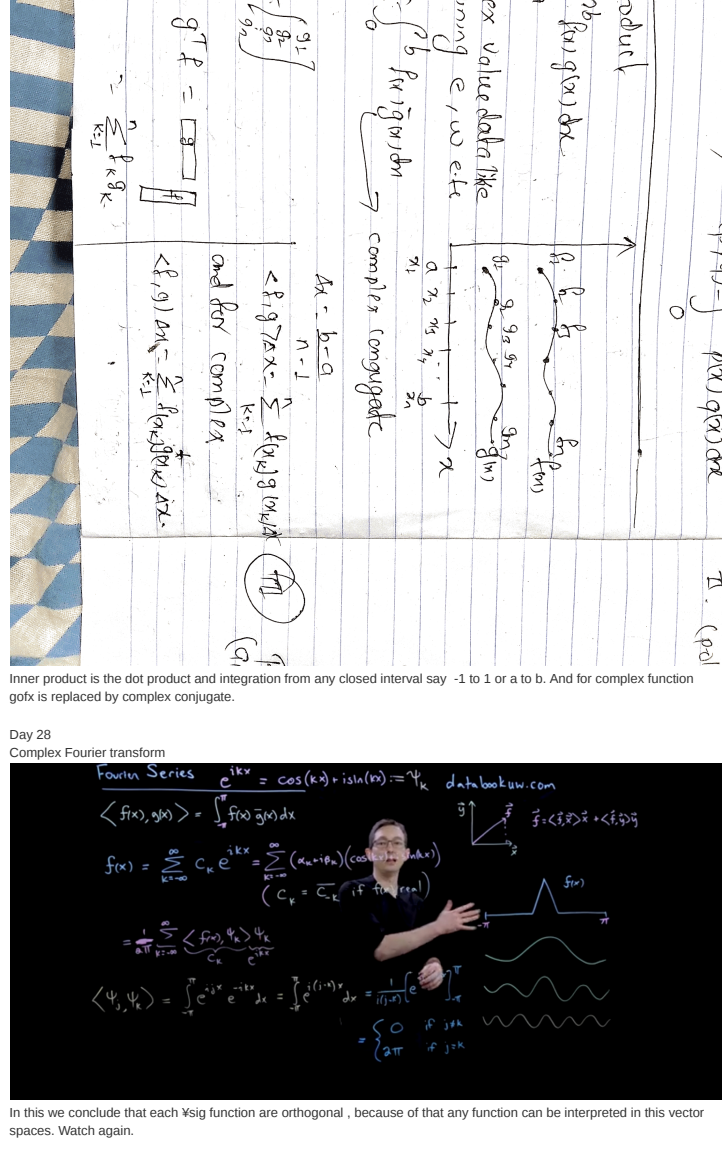
II. (polynomials)

$\langle p,q\rangle = \int_{-1}^{1} p(x)q(x)\,dx$

$\langle x,x\rangle = \int_{-1}^{1} x^2\,dx = \frac{x^3}{3}\Big|_{-1}^{1} = \frac{2}{3}$

$\langle 1,x\rangle = \int_{-1}^{1} x\,dx = 0$ (orthogonal) (Definition of orthogonal)

$\left\langle \sqrt{\tfrac{3}{2}}\,x, \sqrt{\tfrac{3}{2}}\,x\right\rangle = \left(\sqrt{\tfrac{3}{2}}\right)^2\langle x,x\rangle = 1$ (Unit vector)



Inner product

$\langle f(x), g(x)\rangle = \int_a^b f(x)g(x)\,dx$

And for complex value data like function containing e, ω etc

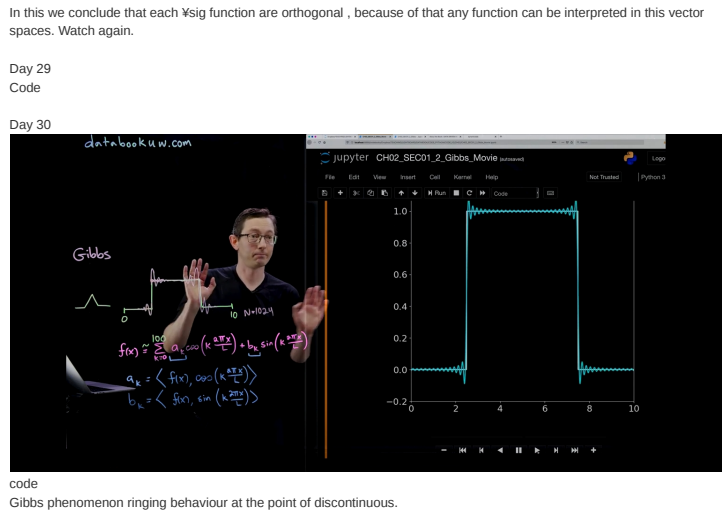
$\langle f(x), g(x)\rangle = \int_a^b f(x)\bar{g}(x)\,dx$

→ complex conjugate

$\langle f,g\rangle = g^T f$

Inner product is the dot product and integration from any closed interval say -1 to 1 or a to b. And for complex function gofx is replaced by complex conjugate.

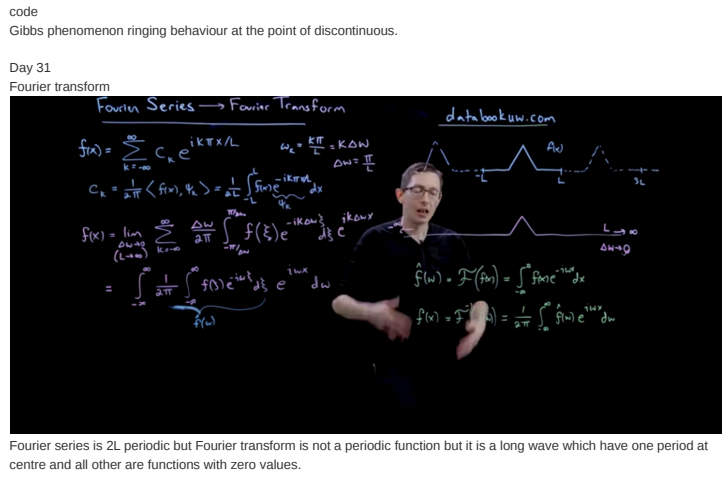
Day 28
Complex Fourier transform


Fourier Series
e^{ikx} = cos(kx) + i sin(kx)
databookuw.com

In this we conclude that each ¥sig function are orthogonal , because of that any function can be interpreted in this vector spaces. Watch again.

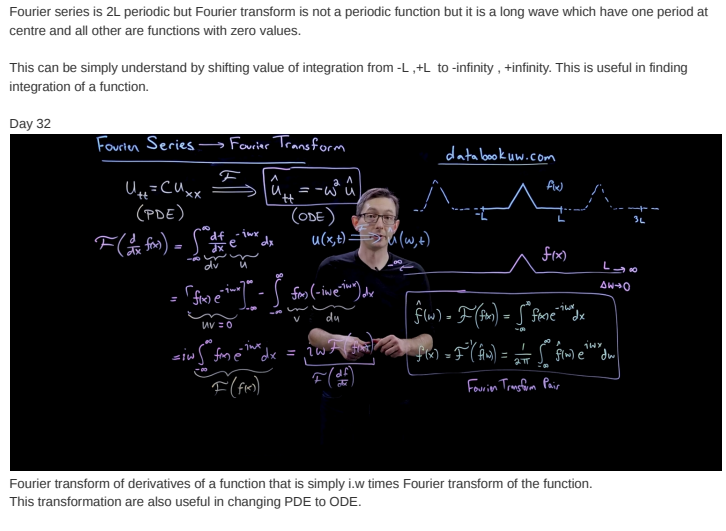
Day 29
Code

Day 30


databookuw.com
Gibbs
jupyter CH02_SEC01_2_Gibbs_Movie

code
Gibbs phenomenon ringing behaviour at the point of discontinuous.

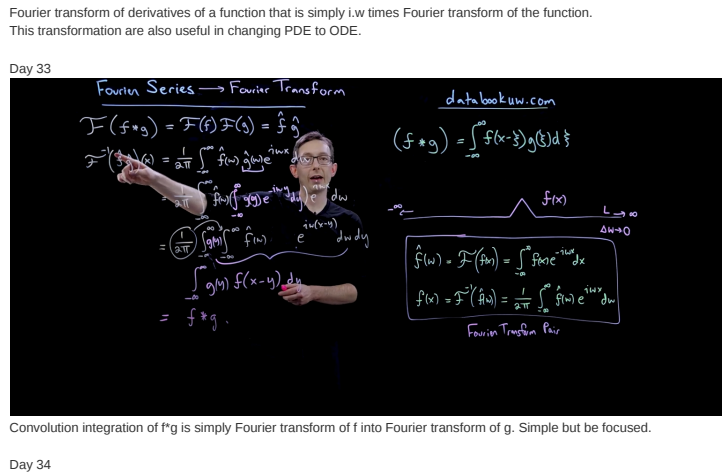
Day 31
Fourier transform


Fourier Series → Fourier Transform
databookuw.com

Fourier series is 2L periodic but Fourier transform is not a periodic function but it is a long wave which have one period at centre and all other are functions with zero values.

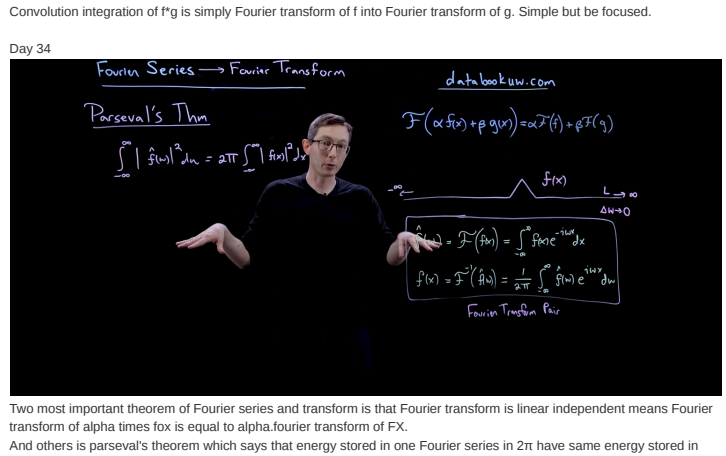
This can be simply understand by shifting value of integration from -L ,+L to -infinity , +infinity. This is useful in finding integration of a function.

Day 32


Fourier Series → Fourier Transform
databookuw.com
(PDE)
(ODE)
Fourier Transform Pair

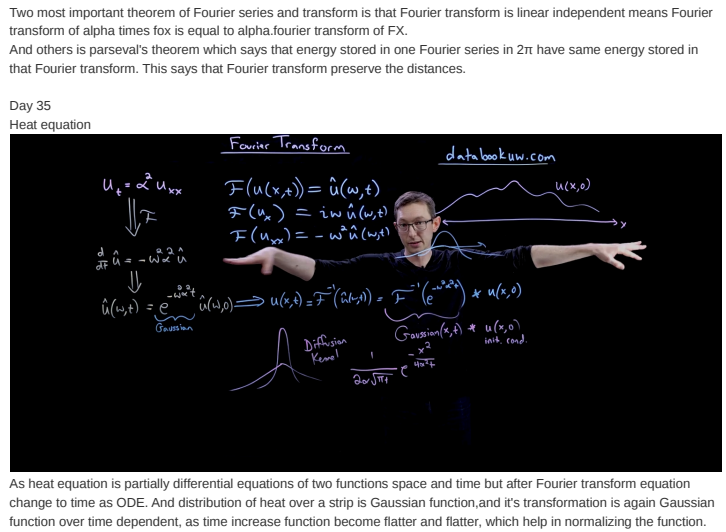
Fourier transform of derivatives of a function that is simply i.w times Fourier transform of the function.
This transformation are also useful in changing PDE to ODE.

Day 33


Fourier Series → Fourier Transform
databookuw.com
Fourier Transform Pair

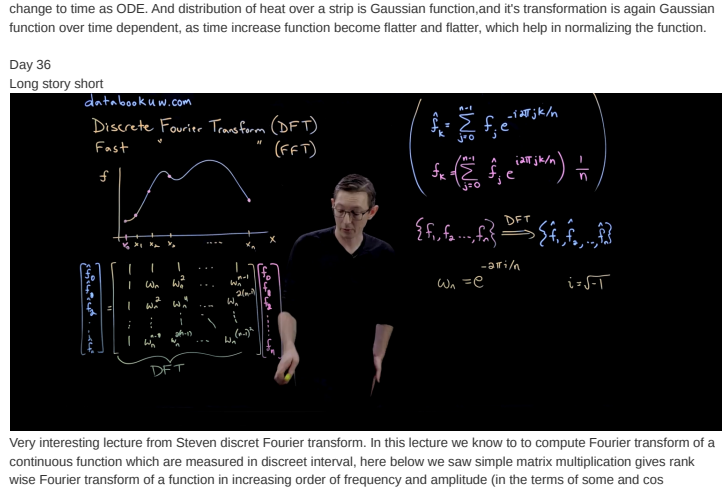
Convolution integration of f*g is simply Fourier transform of f into Fourier transform of g. Simple but be focused.

Day 34


Fourier Series → Fourier Transform
databookuw.com
Parseval's Thm
Fourier Transform Pair

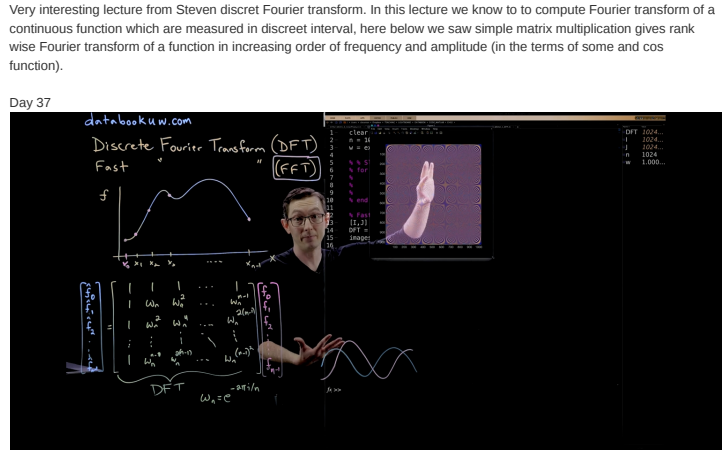
Two most important theorem of Fourier series and transform is that Fourier transform is linear independent means Fourier transform of alpha times fox is equal to alpha.fourier transform of FX.
And others is parseval's theorem which says that energy stored in one Fourier series in 2π have same energy stored in that Fourier transform. This says that Fourier transform preserve the distances.

Day 35
Heat equation


Fourier Transform
databookuw.com
Gaussian
Diffusion Kernel

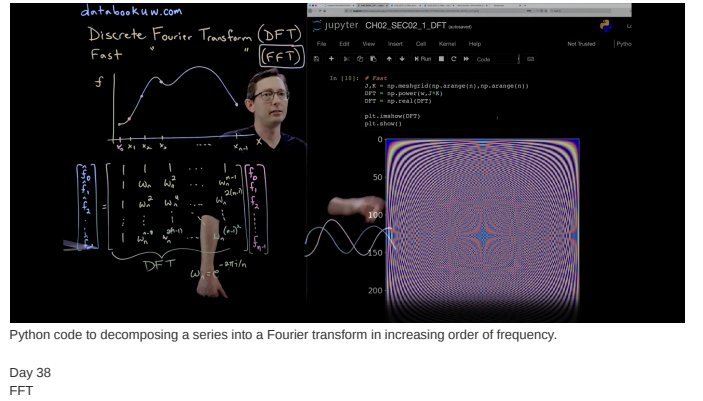
As heat equation is partially differential equations of two functions space and time but after Fourier transform equation change to time as ODE. And distribution of heat over a strip is Gaussian function,and it's transformation is again Gaussian function over time dependent, as time increase function become flatter and flatter, which help in normalizing the function.

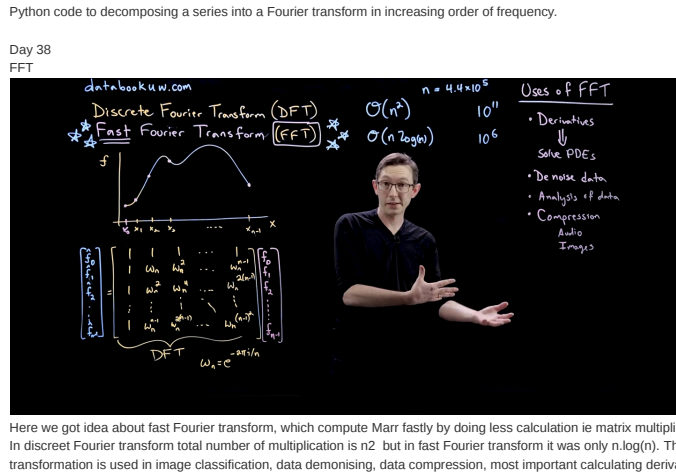
Day 36
Long story short


databookuw.com
Discrete Fourier Transform (DFT)
Fast (FFT)
DFT

Very interesting lecture from Steven discret Fourier transform. In this lecture we know to to compute Fourier transform of a continuous function which are measured in discreet interval, here below we saw simple matrix multiplication gives rank wise Fourier transform of a function in increasing order of frequency and amplitude (in the terms of some and cos function).

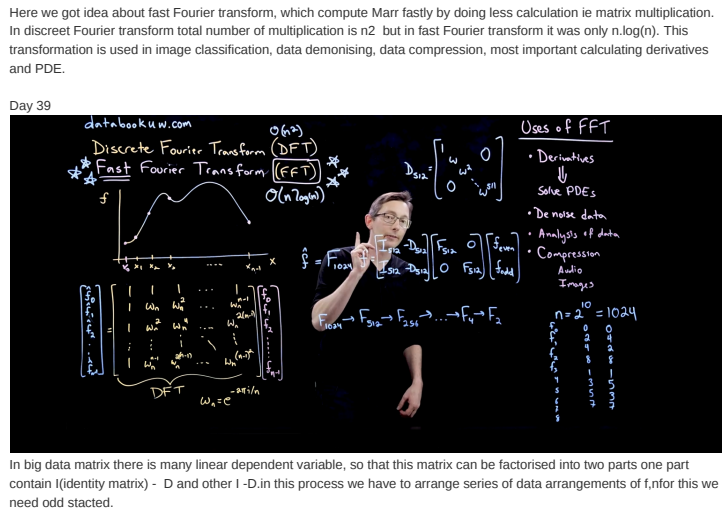
Day 37


databookuw.com
Discrete Fourier Transform (DFT)
Fast (FFT)


databookuw.com
Discrete Fourier Transform (DFT)
Fast (FFT)
jupyter CH02_SEC02_1_DFT

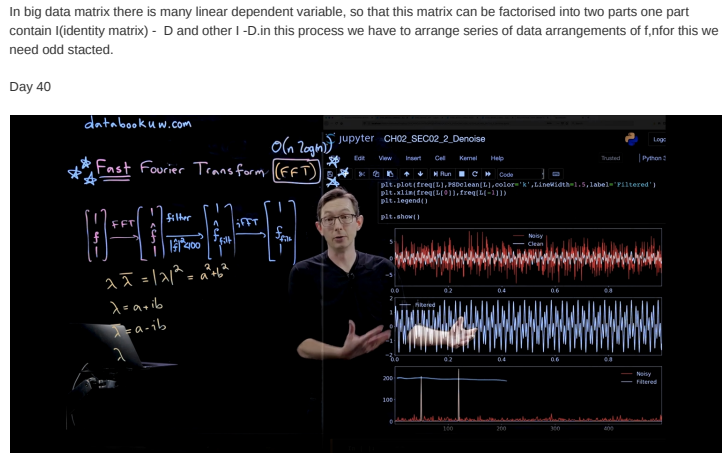
Python code to decomposing a series into a Fourier transform in increasing order of frequency.

Day 38
FFT


databookuw.com
Discrete Fourier Transform (DFT)
Fast Fourier Transform (FFT)
O(n²)
O(n log n)
Uses of FFT
• Derivatives & Solve PDEs
• De-noise data
• Analysis of data
• Compression Audio Images

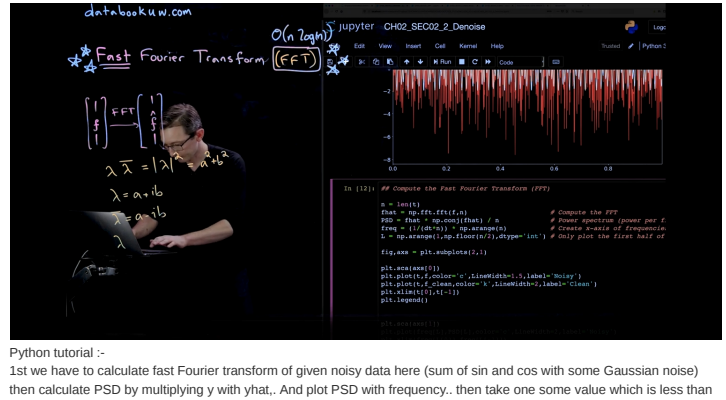
Here we got idea about fast Fourier transform, which compute Marr fastly by doing less calculation ie matrix multiplication. In discreet Fourier transform total number of multiplication is n2 but in fast Fourier transform it was only n.log(n). This transformation is used in image classification, data demonising, data compression, most important calculating derivatives and PDE.

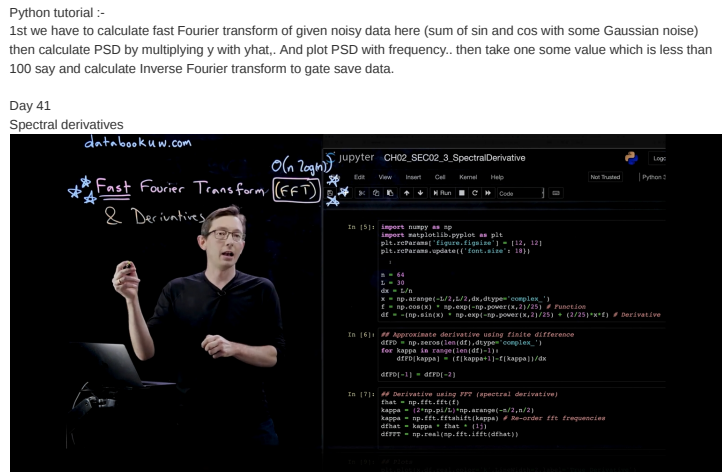
Day 39


databookuw.com
Discrete Fourier Transform (DFT)
Fast Fourier Transform (FFT)
Uses of FFT
n = 2^10 = 1024

In big data matrix there is many linear dependent variable, so that this matrix can be factorised into two parts one part contain I(identity matrix) - D and other I -D.in this process we have to arrange series of data arrangements of f,nfor this we need odd stacted.

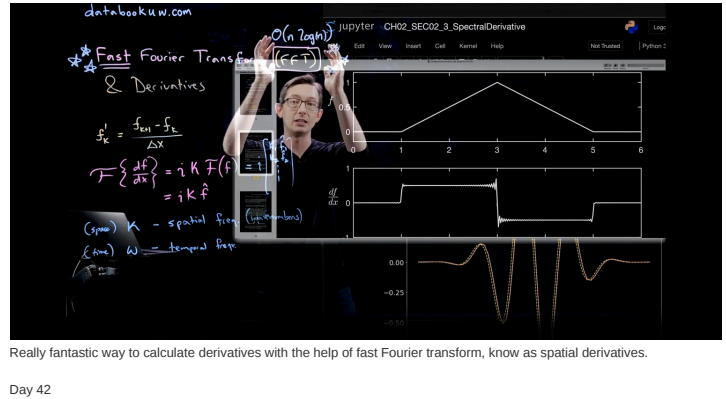
Day 40

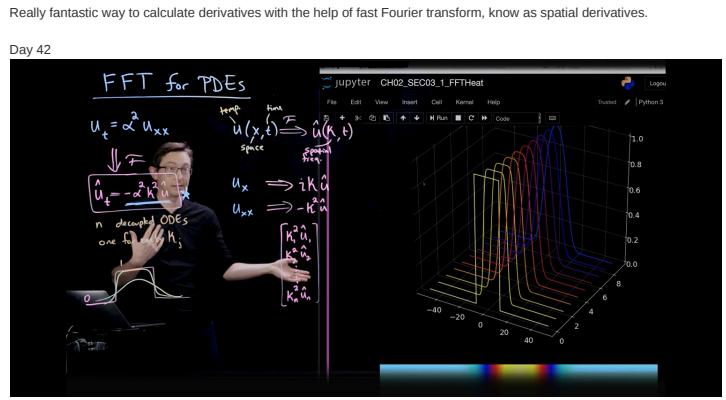

databookuw.com
Fast Fourier Transform (FFT)
O(n log n)
jupyter CH02_SEC02_2_Denoise


databookuw.com
Fast Fourier Transform (FFT)
O(n log n)
jupyter CH02_SEC02_2_Denoise

Python tutorial :-
1st we have to calculate fast Fourier transform of given noisy data here (sum of sin and cos with some Gaussian noise) then calculate PSD by multiplying y with yhat,. And plot PSD with frequency.. then take one some value which is less than 100 say and calculate Inverse Fourier transform to gate save data.

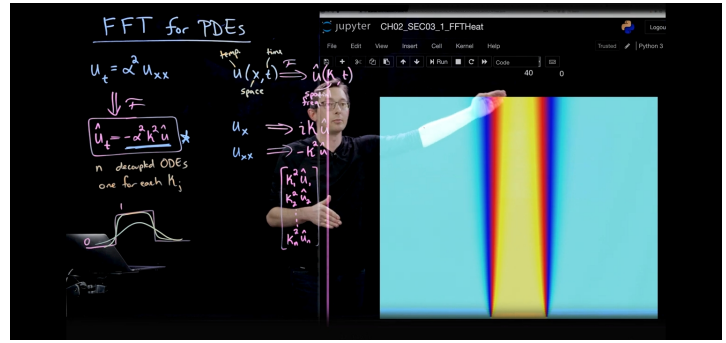
Day 41
Spectral derivatives

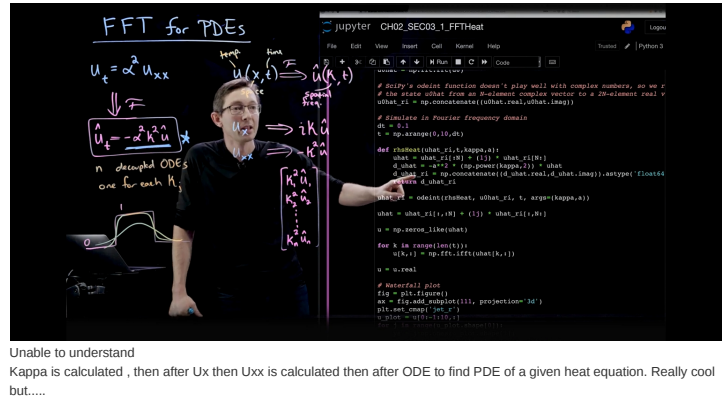

databookuw.com
Fast Fourier Transform (FFT)
& Derivatives
jupyter CH02_SEC02_3_SpectralDerivative

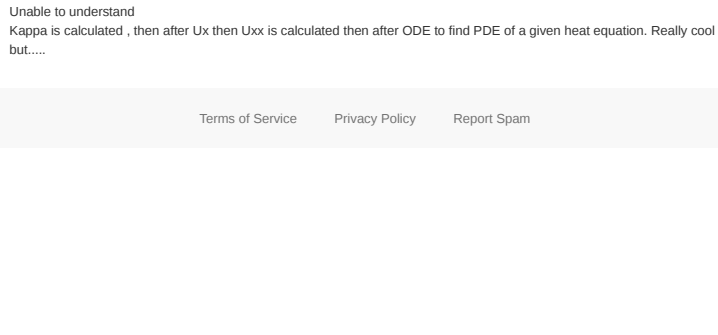

databookuw.com
Fast Fourier Transform (FFT)
& Derivatives
κ - spatial freq
ω - temporal freq

Really fantastic way to calculate derivatives with the help of fast Fourier transform, know as spatial derivatives.

Day 42


FFT for PDEs
jupyter CH02_SEC03_1_FFTHeat
n decoupled ODEs one for each κ


FFT for PDEs
jupyter CH02_SEC03_1_FFTHeat


FFT for PDEs
jupyter CH02_SEC03_1_FFTHeat

Unable to understand
Kappa is calculated , then after Ux then Uxx is calculated then after ODE to find PDE of a given heat equation. Really cool but.....