Lecture 5- Eigenvectors of Circulant Matrices :Fourier Matrix

Agenda

- Circulant Matrices (Image Processing in ML)

- Discrete Fourier Transform : Fourier Matrix

Source: Section IV.2, IV.5 in Linear Algebra and Learning from Data (2019) by **Gilbert Strang**

Circulant Matrices (Image Processing in ML)

Circulant matrices are NxN matrices. Key point, C are defined by not n^2 entries, only n

$$C = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ N-1 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & N-1 \end{array} \\ \begin{pmatrix} c_0 & c_1 & c_2 & \\ & \ddots & \ddots & c_2 \\ c_2 & & \ddots & c_1 \\ c_1 & c_2 & & c_0 \end{pmatrix} \end{array}$$

C gets completed cyclically. Take ML problem having N samples. Let an image with pixels 1,000 by 1,000 pixel- i.e., million pixels, if it is in color, it is 3 million "features"

Say 3 million features, those matrices of size like 3 million times 3 million. Can't use gradient descent to optimize many weights

Matrices in deep learning don't depend on Circulant matrices -- they're like circulant matrices, C. They might not loop around as circulant matrices have this cyclic feature

$$T = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ N-1 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & N-1 \end{array} \\ \begin{pmatrix} t_0 & t_1 & t_2 & \\ t_{-1} & \ddots & \ddots & t_2 \\ t_{-2} & \ddots & \ddots & t_1 \\ & t_{-2} & t_{-1} & t_0 \end{pmatrix} \end{array}$$

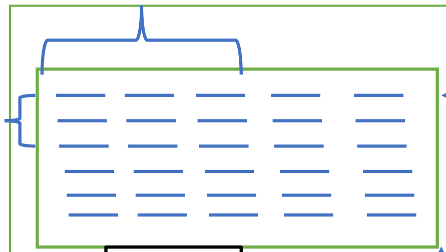
Linear shift invariant (LSI), linear time invariant (LST), it is non-cyclic
Maths people call it a Toeplitz matrix. Used the letter t

In engineering it is a filter or a convolution or a constant diagonal matrix.
These come up in ML and with image processing

Circulant Matrices with convolution

C is multiplied by a vector v, result is the cyclic convolution $Cv=c\otimes v$
Similarly, T multiplied by v, the result is the non-cyclic $Tv=t * v$

Question: At one point in an image what you're going to do at the other points, you're not going to figure out weights for each little pixel in the image



Image

Max pooling operation makes matrices smaller. Let us reduce it

Let us consider an image; In the picture below, 9 pixel marked in image, take 3 at a time.
And replace that 9 pixels by 1 pixel. Max of those 9 numbers
That would be a very simple operation that just reduces the dimension, make it smaller.

Reduce the dimension

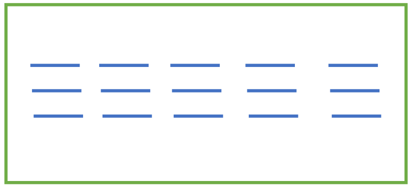
that's a way to make an image 4 times or 9 times or 64 times smaller
But the convolution part now-- so, that's not involving convolution. That's a different operation here
Where do circulants or convolution or Toeplitz matrices or filters come into it?

Circulant Matrices with convolution

Where do circulants or convolution or Toeplitz matrices or filters come into it?
Suppose, a very big system with n^2 pixels, n^2 features for each sample
Operate on that by matrices, as usual

Whole idea is-- on an image use a convolution

Multiply it by weights. What kind of a job would a filter do?



Simpler Image

$$T = \begin{array}{c|cccc} & 1 & 2 & 3 & N-1 \\ \hline 1 & 1/2 & 1/2 & & \\ 2 & & 1/2 & 1/2 & \\ 3 & & & \ddots & \ddots \\ N-1 & & & & \end{array}$$

Job of a filter: kill noise / high frequencies
Let us put some numbers, say ½, ½, .. (low pass filter's values)
averaging each pixel with its neighbor just to take out some of the high frequencies; low frequencies constant

Use a constant diagonal matrix, a shift invariant matrix and not an arbitrary matrix

Only to compute n weights and not n^2- that's the point

That's one reason for convolution and circulant matrices

Circulant Matrices with convolution

Cyclic convolution. So, this would be cyclic because of the looping around stuff
Started with this permutation matrix. And the permutation matrix (P) has c_0 equals 0, c_1 equal 1, and the rest of the c's are 0

Cyclic Shift

$$P=\begin{array}{cc} & \begin{array}{cccc}1 & 2 & 3 & N-1\end{array}\\ \begin{array}{c}1\\2\\3\\N-1\end{array} & \begin{pmatrix} 0 & 1 & 0 & 0\\ & \ddots & \ddots & 0\\ 0 & & \ddots & 1\\ 1 & 0 & & 0\end{pmatrix}\end{array}$$

Look at singular values and then we'll see why it's eigenvalues we want. What are the singular values of a permutation matrix (P). They are all 1

P is orthogonal matrix so the SVD of the matrix just has the permutation and then the identity (I) is there for $\boldsymbol{\Sigma}$. So, $\boldsymbol{\Sigma}$ is I for P. So. $P^TP=I$

Singular values will be the eigenvalues of the identity. And they're all just 1's

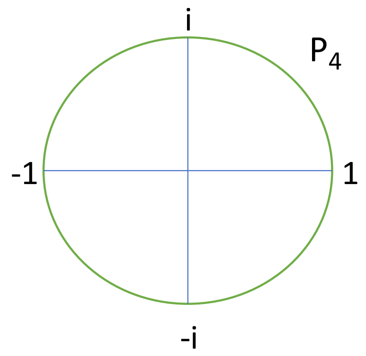
Eigenvalues of P, that's what we want to find, so let's do that

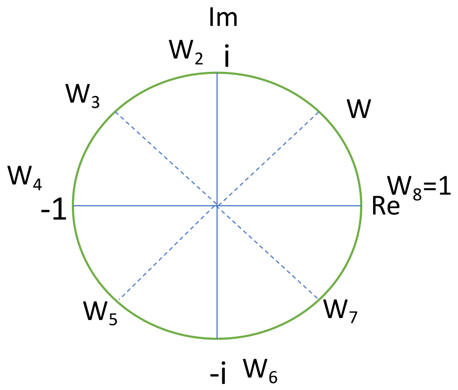
Circulant Matrices with convolution

Eigenvalues of P= P-λI= $\begin{pmatrix} -\lambda & 0 & 0 & 0 \\ 0 & -\lambda & 0 & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & 0 & -\lambda \end{pmatrix}$

Put -λ on diagonal,
and rest is 0

det (P-λI)=0$\Rightarrow$ λ^4-1=0
$\Rightarrow$ λ=1,-1, i, -i





The roots of 1 are equally spaced around the circle.
So, Fourier has come in and Fourier is going to be here and it'll be in the eigenvectors

$w=e^{2\pi i/8}= e^{2\pi i/N}$ (for NxN matrix)

Eigenvalues for the P(8x8)
P_8=I so λ^8=1

These eigenvectors of an orthogonal matrix are orthogonal just like symmetric matrices. Call them q

Matrices of orthogonal evectors : $(q_i, q_j)=1$, i=j and $(q_i, q_j)=0$, $i\neq j$

Circulant Matrices with convolution

Matrices with orthogonal eigenvectors are Symmetric Matrices

Symmetric matrices have orthogonal eigenvectors and eigenvalues will be real, i.e., λ (real)

Other kind of matrices have orthogonal eigenvectors they might be complex, and eigenvalues are complex

Other family of matrices are diagonal and eigen vectors are identity (I) matrix, diagonals are 1's

Orthogonal matrices, special of its eigen values, magnitude $|\lambda|=1$ i.e., is $||Qx||=||x||$ for all x. Also, $||Qx||=||\lambda x||$

So, from above, Symmetric matrix $A^T=A$, diagonal matrix $AI=\Lambda I$ and Orthogonal matrix $A^T=A^{-1}$

Anti-symmetric matrix $A^T=-A$. Take an example $A-\lambda I=\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2+1=0$ So, $\lambda=i,-i$

That's a rotation matrix. It's a rotation through 90 degrees. So there could not be a real eigenvalue.

So, $A^T=-A$ (λ imaginary)

Circulant Matrices with convolution

Finally, the name of the whole family of matrices includes all of those and more of matrices with orthogonal eigenvectors

Take M for matrix. $M=Q\Lambda Q^T$ -- So, it has orthogonal eigenvectors. So it's Q times diagonal times Q transpose

A matrix of that form is a normal matrix. Normal. How do you recognize a normal matrix? Quick test $M^TM=MM^T$

Matrices might be real but the eigenvectors are not going to be real and the eigenvalues are not going to be real

How do we get out of the limitation to real? if M is a complex matrix instead of a real matrix? Then whenever you transpose it you should take its complex conjugate: $\overline{M}^TM=M\overline{M}^T$

$$C=\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ N-1 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & N-1 \end{array} \\ \begin{pmatrix} C_0 & C_1 & C_2 & \\ & & & C_2 \\ C_2 & & & C_1 \\ C_1 & C_2 & & C_0 \end{pmatrix} \end{array}$$

Permutation P is orthogonal so its eigenvectors. Circulant matrix C is normal matrix. Because $C_1C_2= C_2C_1$ Two circulant matrices commute

eigenvectors of P will also be eigenvectors of C because $C=C_0I+C_1P+C_2P^2+\ldots+C_{n-1}P^{n-1}$

Eigenvectors of P is the eigenvectors of C. And these eigenvectors connection to Fourier

Circulant Matrices with convolution

$$P=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}\quad \underset{\lambda=1}{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}}\quad \underset{\lambda=-1}{\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}}\quad \underset{\lambda=i}{\begin{pmatrix} 1 \\ i \\ i^2 \\ i^3 \end{pmatrix}}\quad \underset{\lambda=-i}{\begin{pmatrix} 1 \\ -i \\ (-i)^2 \\ (-i)^3 \end{pmatrix}}$$

1, 1, 1, 1. make it into a vector. And the eigenvector for λ =- 1 this shift to change every sign. eigenvalue i . Start it with 1 and permutation, then i, i^2, i^3 there with - i, the vector 1, - i, - i^2, -i^3.

Eigenvector matrix for all circulants of size C. They all have the same eigenvectors, including P.

All circulants C_N of size N including P_N of size N

So, $w= e^{2\pi i/N}$ =>$w^8=1$ with eight eigen vectors and eigen values

So, that is an orthogonal matrix. It has orthogonal columns but it doesn't have orthonormal columns.
What's the length of that column vector? That is √8.
divide out √8 to make the columns orthonormal. F= √8Q

F_8 (8x8) Fourier Matrix =eigen vector matrix
$W^8=1 \rightarrow w^{48}=1$ and $w^{49}=w$

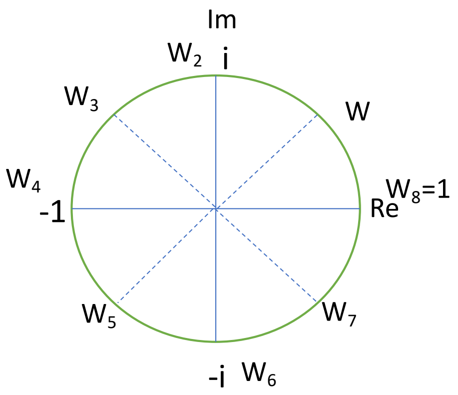
$$F_8=\begin{array}{c|cccc} & 1 & 2 & 3\ \ldots & 8 \\ \hline 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & w & w^2 & w^7 \\ 3 & 1 & w^2 & w^4 & w^{14} \\ 4 & 1 & w^3 & w^6 & w^{21} \\ 5 & 1 & w^4 & w^8 & w^{28} \\ 6 & 1 & w^5 & w^{10} & w^{35} \\ 7 & 1 & w^6 & w^{12} & w^{42} \\ 8 & 1 & w^7 & w^{14} & w^{49} \end{array}$$

Circulant Matrices with convolution

Fourier matrix is a normal matrix. Normal matrices have orthogonal--How do we know it's a normal matrix

Do test: P → Permutation matrix and $P^TP=PP^T$ and they commute i.e.,$P_1P_2=P_2P_1$, so it's a normal matrix

So that dot product of Column 1 and Column 2 is 0, i.e., $1+w+w^2+w^3+w^4+w^5+w^6+w^7$



There's symmetry. When add w to the 0, or w to the eighth, or w to the 0. when add 1 and minus 1, get 0. When add w^5 to w to get 0. When add these by pairs to get zeros