Reminder: Eigenvalues and eigenvectors

$A \cdot x = \lambda \cdot x \leftarrow$ corresponding eigenvector

$\uparrow$ eigenvalue of A

Assume A is $n \times n$ and has n eigenvectors (a basis):

$$A \cdot x_i = \lambda_i \cdot x_i, \qquad i = 0..n-1$$

Set $T = (x_0 | \dots | x_{n-1})$

$$\Rightarrow \quad A \cdot T = (\lambda_0 x_0 | \dots | \lambda_{n-1} x_{n-1})$$

$$= T \cdot \operatorname*{diag}_{i=0..n-1} (\lambda_i)$$

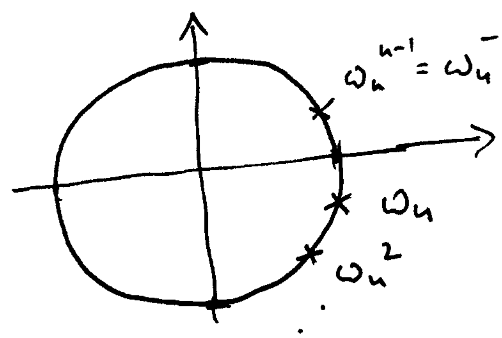
$$\Rightarrow \quad T^{-1} A T = \operatorname{diag}(\lambda_i)$$

T diagonalizes A.

DFT

1.) Definition $\quad DFT_n = [\omega_n^{kl}]_{k,l=0..n-1}, \quad \omega_n = e^{-\frac{2\pi j}{n}}, \quad j = \sqrt{-1}$

complex plane:



- roots of unity
- zeros of $x^n - 1$

$$\omega_n = e^{-\frac{2\pi j}{n}} = \cos \frac{2\pi}{n} - j \sin \frac{2\pi}{n}$$

$$\omega_n^k = e^{-\frac{2\pi k j}{n}} = \cos \frac{2\pi k}{n} - j \sin \frac{2\pi k}{n}$$

[Note: In math, $\omega_n = e^{\frac{2\pi j}{n}}$ is chosen]

2.) Interpretation of the DFT

a.) Let $x = (p_0, \dots, p_{n-1})^T$, the kth output of the DFT is

$$y_k = \sum_{l=0}^{n-1} p_l \omega_n^{kl} = \sum_{l=0}^{n-1} p_l (\omega_n^k)^l$$

is the evaluation of $p(x) = \sum p_l x^l$ at ω_n^k

$$\Rightarrow \quad DFT_n \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_{n-1} \end{pmatrix} = \begin{pmatrix} p(\omega_n^0) \\ p(\omega_n^1) \\ \vdots \\ p(\omega_n^{n-1}) \end{pmatrix}$$

The DFT evaluates a polynomial of degree $n-1$ at all n nth roots of unity.

[This implies that this simultaneous evaluation can be done in $O(n \log(n))$ using FFTs.]

b.) $C_n = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 1 & & & 0 \end{pmatrix}$ is called cyclic shift

[Note: C_n is a permutation matrix (every row and column exactly one 1). This implies $C_n^T = C_n^{-1}$.]

Pick the lth column of DFT_n:

$$C_n \cdot \begin{pmatrix} (\omega_n^l)^0 \\ (\omega_n^l)^1 \\ \vdots \\ (\omega_n^l)^{n-1} \end{pmatrix} = \begin{pmatrix} (\omega_n^l)^1 \\ \vdots \\ (\omega_n^l)^{n-1} \\ (\omega_n^l)^0 \end{pmatrix} = \omega_n^l \begin{pmatrix} (\omega_n^l)^0 \\ \vdots \\ \vdots \\ (\omega_n^l)^{n-1} \end{pmatrix}$$

$\Rightarrow$ Every column of DFT_n is an eigenvector of C_n. The lth column has eigenvalue ω_n^l.

$$\Rightarrow \; DFT_n^{-1} C_n \, DFT_n = \operatorname*{diag}_{\ell=0..n-1}(\omega_n^{\ell})$$

or $$C_n \, DFT_n = DFT_n \cdot \operatorname{diag}(\omega_n^{\ell})$$

means: first scaling and then DFT

= first DFT and then cyclic shift

Transpose to get:

$$DFT_n^{T} \cdot C_n^{T} = \operatorname{diag}(\omega_n^{\ell})^{T} \cdot DFT_n^{T}$$

$$\Leftrightarrow \; \operatorname{diag}(\omega_n^{\ell}) \cdot DFT_n = DFT_n \cdot C_n^{-1}$$

means: first cyclic shift (other direction) and then DFT

= first DFT and then scaling

General diagonalization properties:

$$DFT_n^{-1} \cdot C_n \cdot DFT_n = \operatorname{diag}(\omega_n^{\ell})$$

$$\Rightarrow \; DFT_n^{-1} \cdot C_n^{2} \cdot DFT_n = \operatorname{diag}((\omega_n^{\ell})^{2}) \qquad C_n^2 = \begin{pmatrix} 0 & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 1 & & & & 0 \\ 0 & 1 & & & 0 \end{pmatrix}$$

$$\Rightarrow \; DFT_n^{-1} \cdot C_n^{k} \cdot DFT_n = \operatorname{diag}((\omega_n^{\ell})^{k}) \quad \text{for all } k$$

$$\Rightarrow \; DFT_n^{-1} \cdot \underbrace{\sum_{k=0}^{n-1} p_k C_n^{k}}_{\text{is called circulant matrix}} \cdot DFT_n = \operatorname*{diag}_{\ell=0..n-1}\Big(\sum_{k=0}^{n-1} p_k (\omega_n^{\ell})^{k}\Big)$$

$$\operatorname{Circ}(p_0, \dots, p_{n-1}) = \begin{pmatrix} p_0 & p_1 & p_2 & \cdots & \cdots \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & p_2 \\ p_2 & \cdots & & \ddots & p_1 \\ p_1 & p_2 & & \cdots & p_0 \end{pmatrix}$$

The DFT diagonalizes all circulant matrices (of the same size) and only those.

Equivalent formulation:

Let $p(x) = \sum_{k=0}^{n-1} p_k x^k$

$\Rightarrow$ DFT_n diagonalizes $p(C_n)$, namely

$$DFT_n^{-1} \cdot p(C_n) \cdot DFT_n = \operatorname{diag}_{l=0}^{n-1} (p(\omega_n^l))$$

$\Rightarrow$ $\operatorname{Circ}(p_0, \dots, p_{n-1}) = p(C_n) = DFT_n \cdot \operatorname{diag}(p(\omega_n^l)) \cdot DFT_n^{-1}$

$\Rightarrow$ <u>Multiplying by a circulant is $O(n \log(n))$</u>

3.) Toeplitz Matrices:

$$\underset{(n\times n)}{A} = \begin{pmatrix} a_0 & a_1 & \cdots & \cdot \\ a_{-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ \cdot & \cdots & a_{-1} & a_0 \end{pmatrix}$$

Determined by $2n-1$ parameters: $a_{-(n-1)}, \dots, a_0, \dots, a_{n-1}$

In words: Those matrices that are constant along all diagonals

Every Toeplitz matrix can be embedded into a circulant of twice the size:

A Toeplitz : $C = \begin{pmatrix} A & B \\ B & A \end{pmatrix}$ circulant

Example: $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$, $C = \left(\begin{array}{cc|cc} 2 & 3 & x & 1 \\ 1 & 2 & 3 & x \\ \hline x & 1 & 2 & 3 \\ 3 & x & 1 & 2 \end{array}\right)$ x: does not matter

$$y = Ax \quad \longleftrightarrow \quad \begin{pmatrix} y \\ \hline * \end{pmatrix} = \begin{pmatrix} A & B \\ B & A \end{pmatrix} \begin{pmatrix} x \\ \hline 0 \end{pmatrix}$$

$\Rightarrow$ <u>Multiplying by a Toeplitz matrix $(n\times n)$ is $O(n \log(n))$</u>

DFT Algorithms

1.) Structured Matrices

$I_n = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$, $\mathrm{DFT}_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, $\mathrm{diag}(a_0, \dots, a_{n-1}) = \begin{pmatrix} a_0 & & & 0 \\ & a_1 & & \\ & & \ddots & \\ 0 & & & a_{n-1} \end{pmatrix}$

$A \oplus B = \begin{pmatrix} A & \\ & B \end{pmatrix}$,

$A \otimes B = [a_{k,\ell} \cdot B]$, where $A = [a_{k,\ell}]_{k,\ell = 0 \dots n-1}$

In particular:

$I_n \otimes A = \begin{pmatrix} A & & & \\ & A & & \\ & & \ddots & \\ & & & A \end{pmatrix}$ "n times A at stride 1"

$A \otimes I_n = [a_{k,\ell} \cdot I_n] = \begin{pmatrix} \ddots & \ddots & \ddots & \ddots \\ \ddots & \ddots & \ddots & \ddots \\ & \dots & & \end{pmatrix}$ "n times A at stride n"

Where does it occur?

$X = \underbrace{\begin{bmatrix} \text{row}_1 \\ \text{row}_2 \\ \vdots \end{bmatrix}}_{n} \Big\} n$ in memory: $\mathrm{vec}(X) = (\text{row}_1, \text{row}_2, \dots)$

$$X \cdot A \quad \leftrightarrow \quad (I_n \otimes A^T)\,\mathrm{vec}(X)$$

$$A \cdot X \quad \leftrightarrow \quad (A \otimes I_n)\,\mathrm{vec}(X)$$

$$\begin{aligned} A \cdot X \cdot A \quad \leftrightarrow \quad & (A \otimes I_n)(I_n \otimes A^T)\,\mathrm{vec}(X) \\ = \ & (I_n \otimes A^T)(A \otimes I_n)\,\mathrm{vec}(X) \\ = \ & (A \otimes A^T)\,\mathrm{vec}(X) \end{aligned}$$

~~2D version of transform A~~

Property: $(A \otimes B)(C \otimes D) = (AC \otimes BD)$ if matrix sizes are compatible

$A \otimes A$: 2D version of transform A