SVD – Singular value decomposition.

→ decomposes matrix A into r rank 1 matrices

$$A = \sum \sigma_i u_i v_i^T \quad ; A \in \mathbb{R}^{m \times n}$$

where u_i and v_i are orthonormal.

we also have $\sigma_1 \geq \sigma_2 \dots \geq \sigma_r > 0$

→ SVD can also be used to make the best $\hat{r} < r$ rank approximation of A.

i.e. $A \approx \sum_{\hat{r}} \sigma_i u_i v_i^T$

→ Matrix form $A = U_1 S (V_1^T)$

Compact form → $= \underbrace{[u_1\ u_2 \dots u_r]}_{m \times r} \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}}_{r \times r} \underbrace{\begin{bmatrix} - v_1^T - \\ - v_2^T - \\ \vdots \\ - v_r^T - \end{bmatrix}}_{r \times n}$

Full-matrix form → $= \underbrace{[U_1\ U_2]}_{m \times m} \underbrace{\begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}}_{(m \times n)} \underbrace{\begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}}_{n \times n}$

note U_1 spans the col space of A, and V_1 spans the row space of A.

→ How SVD? A^TA or AA^T

(i) eigenvalues of A^TA, $\lambda_i = \sigma_i^2$

(ii) eigenvector of A^TA, v_i from spectral thm.

(iii) $u_i = \frac{1}{\sigma_i} A\vec{v}_i$

note: A^TA is symmetric ⇒ $\{v_i\}$ are orthonormal

→ Q: S.T. $\{u_i\}$ are orthonormal.

→ TODO: review computing SVD for AA^T @ home

when do you use AA^T vs. A^TA ? $A \in \mathbb{R}^{m \times n}$

→ Geometric interpretation: $A^TA = n \times n$; $AA^T = m \times m$

- U and V only rotate our input vector
- S will scale our vectors.

from this interpretation: $\|A\vec{x}\| \leq \sigma_1 \|\vec{x}\|$, $\|\vec{x}\| = 1$

when is it equal? $x = ?$

PCA – Principal component analysis:

→ Given some attributes, PCA finds the most relevant attributes which separate the data most efficiently

→ Maximize variance, minimize projection error.

→ rows contain data points and columns contain attributes/features.

$A \in \mathbb{R}^{n \times d}$, perform SVD $A = U_1 S V_1^T$

The top k right singular vectors $\{v_i\}_{i=1 \ldots k}$ gives "new" attributes which best separate the data.

→ $XV = USV^TV = US$ give the projections along all the $\{v_i\}$

Interpolation

→ fit a polynomial passing through n points. We will need a polynomial of degree at least $(n-1)$.

$$p(x) = a_0 + a_1 x + a_2 x^2 \dots + a_{n-1} x^{n-1}$$

→ We want

$$p(x_1) = y_1 ; \quad p(x_2) = y_2 ; \quad \vdots \quad p(x_n) = y_n \Rightarrow \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ & \vdots & & & \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Vandermonde matrix (V)

$$\det(V) = \prod_{1 \le i < j \le n} (x_j - x_i)$$

$$= (x_n - x_{n-1})(x_n - x_{n-2}) \dots (x_n - x_1).$$
$$(x_{n-1} - x_{n-2})(x_{n-1} - x_{n-3}) \dots (x_{n-1} - x_1).$$
$$\vdots$$
$$(x_3 - x_2)(x_3 - x_1).$$
$$(x_2 - x_1)$$

$\because x_i \neq x_j$ for $i \neq j$, $\det(V) \neq 0 \Rightarrow V$ is invertible, hence a unique solution exists.

→ Basis functions and kernel interpolation:

define kernel $\Phi(x) \begin{cases} 1 & x = 0 \\ 0 & x = k\Delta \end{cases}$

where Δ is sampling period, $x_{i+1} - x_i = \Delta$

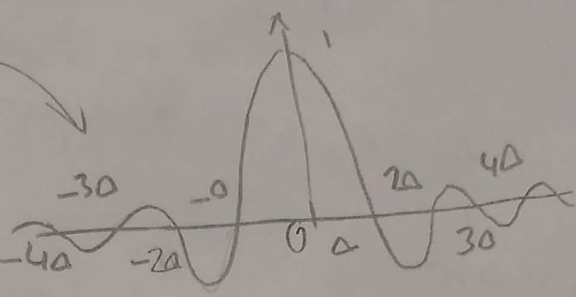
$$\therefore y(x) = \sum_{k=1}^{N} y_k \Phi(x - k\Delta)$$

Ex: linear interpolation.

→ use $\mathrm{sinc}\left(\frac{x}{\Delta}\right)$ as kernel.

$$\mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x} = \begin{cases} \frac{\sin(\pi x)}{\pi x} & x \neq 0 \\ 1 & x = 0 \end{cases}$$

$$\phi(x) = \mathrm{sinc}\left(\frac{x}{\Delta}\right)$$



→ Interpolation w/ sinc( )

$$y(x) = \sum_{k=-\infty}^{\infty} y_k \, \Phi(x - k\Delta)$$

but since has freq. range b/w $[0, \pi]$, hence we need to sample at some rate for recovery of signal.

→ Sampling theorem | Nyquist criterion

$$w_{max} < \frac{\pi}{\Delta} \Rightarrow 2\pi f_{max} < \pi f_s \leftarrow \text{sampling frequency}$$

$\Rightarrow 2 f_{max} < f_s$, i.e. sample more than twice as fast as the highest freq.

- If we don't sample fast enough we have

Aliasing

Discrete time systems:

→ Causality: $y[n_0]$ depends only on $x[n]$ for $-\infty \leq n \leq n_0$

→ Linearity: let $y[n] = F\{x[n]\}$

- Scaling: $F\{a x[n]\} = a F\{x[n]\} = a y[n]$
- Superposition: $F\{x_1[n] + x_2[n]\}$
  $= F\{x_1[n]\} + F\{x_2[n]\}$
  $= y_1[n] + y_2[n]$

→ BIBO stability: If $x[n]$ is bounded, then $y[n]$ is bounded.

→ Time invariance: Shifted input ⇒ shifted output
i.e. $y[n-n_0] = F\{x[n-n_0]\}$.

→ $x[n] \longrightarrow \boxed{h} \rightarrow y[n]$, let h be impulse response

then $y[n] = \sum_{m=-\infty}^{\infty} x[m]\, h[n-m] = x[n] * h[n]$

convolution.

where the impulse response, $h[n]$, is the output of the system when the input is $\delta[n]$.

→ For LTI system:
- → causal iff $h[n] = 0 \;\forall\; n < 0$
- → stable iff $\sum_{n=-\infty}^{\infty} |h[n]| < \infty$ BIBO

DFT – Discrete Fourier Transform:

→ <u>Basis</u>:

- At the most fundamental level, DFT is just a change of basis.
- We choose the basis vectors

$$\rightarrow \underline{\underline{u_k[n]}} = \frac{1}{\sqrt{N}} e^{j\frac{2\pi}{N}kn}$$, where

$N \equiv$ # of samples, ~~for~~ $k, n \in [0, 1 \dots N-1]$

$$\vec{u}_k = [u_k[0] \ u_k[1] \dots \ u_k[N-1]]^T$$

- Expressing our signal $\vec{x} = [x(0) \ x(1) \dots x(N-1)]^T$ using this basis

$$\vec{x} = \sum_{k=0}^{N-1} \underline{\underline{X[k]}} \vec{u}_k$$, and $X[k]$ ~~is~~ is the

k^{th} frequency component.

$$\vec{X} = [X[0] \ X[1] \dots X[N-1]]^T$$

$$\vec{x}(n) \xrightarrow{DFT} \vec{X}(k)$$

- from above defn. $X[k] = \langle \vec{u}_k, \vec{x} \rangle = \vec{u}_k^* \vec{x}$

$$\Rightarrow \begin{bmatrix} X[0] \\ X[1] \\ \vdots \\ X[N-1] \end{bmatrix} = \begin{bmatrix} \vec{u}_0^* \\ \vec{u}_1^* \\ \vdots \\ \vec{u}_{N-1}^* \end{bmatrix} \vec{x} = F^* \vec{x}$$

where $F = \begin{bmatrix} | & | & & | \\ \vec{u}_0 & \vec{u}_1 & \dots & \vec{u}_{N-1} \\ | & | & & | \end{bmatrix}$

ex: DFT of pure cosine:

Let's say we have a signal $x[n] = \cos\left(\frac{2\pi}{N} kn\right)$,

then $x[n] = \dfrac{e^{j\frac{2\pi kn}{N}} + e^{-j\frac{2\pi}{N}kn}}{2}$

$$= \frac{\sqrt{N}}{2} \cdot \left[\frac{1}{\sqrt{N}} e^{j\frac{2\pi}{N}kn} + \frac{1}{\sqrt{N}} e^{-j\frac{2\pi}{N}kn}\right]$$

$$= \frac{\sqrt{N}}{2} \cdot \left[\vec{u}_k[n] + \frac{1}{\sqrt{N}} e^{j\frac{2\pi}{N}(N-k)n}\right]$$

$$= \frac{\sqrt{N}}{2} \cdot \left(\vec{u}_k[n] + \vec{u}_{N-k}[n]\right)$$

$\Rightarrow X[k] = \frac{\sqrt{N}}{2}$; $X[N-k] = \sqrt{N}/2$ //

$$\vec{X} = [\langle \vec{x}, \vec{u}_0 \rangle \ \langle \vec{x}, \vec{u}_1 \rangle \ \cdots \ \langle \vec{x}, \vec{u}_{N-1} \rangle]$$

$$\vec{X} = F^* \vec{x} \qquad F = \begin{bmatrix} | & & | \\ \vec{u}_0 & \cdots & \vec{u}_{N-1} \\ | & & | \end{bmatrix}$$

$$F \vec{X} = \vec{x} \qquad F^{-1} = F^* \qquad F \text{ unitary matrix}$$

(analysis) $$X[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi k n}{N}}$$

$$X[k] = \langle \vec{x}, \vec{u}_k \rangle = \vec{u}_k^* \vec{x}$$

(synthesis)

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X[k] e^{j\frac{2\pi k n}{N}}$$

Properties

① Linearity

$$DFT\{\alpha_1 \vec{x}_1 + \alpha_2 \vec{x}_2\}$$

$$= \alpha_1 DFT\{\vec{x}_1\} + \alpha_2 DFT\{\vec{x}_2\}$$

α_1, α_2 are scalars

ex: DFT of a cosine

$$x[n] = \cos\left(\frac{2\pi n}{N}\right) \quad n = 0, \ldots, N-1$$

$$x[n] = \frac{1}{2}\left(\underline{e^{j\frac{2\pi n}{N}}} + \underline{e^{-j\frac{2\pi n}{N}}}\right)$$

② Parseval's Theorem (Energy Conservation)

$$\underbrace{\sum_{n=0}^{N-1} |x[n]|^2}_{} = \sum_{k=0}^{N-1} |X[k]|^2$$

$$\|x\|^2 = \|X\|^2$$ "DFT preserves norms"

③ Conjugate Symmetry

if $\vec{x} \in \mathbb{R}^N$,

$$\overline{X[k]} = X^*[k] = X[N-k] \quad k = 1, \dots, N-1$$

x is real

$\Rightarrow x[n] = \cancel{x^*}[n]$

$$X = \begin{bmatrix} X[0] \\ X[1] \\ \vdots \\ \vdots \\ X[N-1] \end{bmatrix}$$

complex conjugates

when something is real

$\vec{y} \quad \vec{y} = \overline{\vec{y}}$

Real$\{X[k]\}$ is even-symmetric

Real$\{X[k]\}$ = Real$\{X[N-k]\}$ $\quad a + jb$

Imag$\{X[k]\}$ is odd-symmetric $\quad$ vs

$\mathrm{Im}\{X[k]\} = -\mathrm{Im}\{X[N-k]\}$ $\quad a - jb$

when I take the DFT of a real sequence I only need half the coefficients!

$X[0], \dots X[N/2]$

(4) Modulation and circular shifting

↳ Multiplying by a complex exponential that varies in time-index

$$x[n] \underbrace{e^{j\frac{2\pi k_0 n}{N}}} \underset{DFT}{\Rightarrow} X[\underbrace{\text{mod}_N(k-k_0)}]$$

length 4 $\quad x[n] \rightarrow \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$

$$\underbrace{e^{j\frac{\pi n}{2}}} x[n] \rightarrow \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}}$$

$$e^{j\frac{2\pi k_0 n}{N}} = e^{j\frac{2\pi k_0 n}{4}}$$

$$\underbrace{k_0 = 1}$$

$$X[k] e^{-j\frac{2\pi k n_0}{N}} \underset{IDFT}{\Rightarrow} x[\text{mod}_N(n-n_0)]$$

(5) DFT Diagonalizes Circulant Matrices

$\vec{h} \rightarrow$ length N

$$C = \begin{bmatrix} h[0] & h[N-1] & \cdots & h[1] \\ h[1] & h[0] & & h[2] \\ \vdots & \vdots & & \vdots \\ h[N-1] & h[N-2] & \cdots & h[0] \end{bmatrix}$$

Result: $F^* C F = \sqrt{N} \begin{bmatrix} H[0] & & & 0 \\ & H[1] & & \\ & & \ddots & \\ 0 & & & H[N-1] \end{bmatrix}$

LTI System

$x[n] \rightarrow \boxed{LTI} \rightarrow y[n]$ $h[n]$ impulse response

$$y[n] = x[n] * h[n]$$

x is length N

h is length M

1) zero-pad $\vec{x}, \vec{h}$ to have lengths $N+M-1$

$\vec{x}_{zp}, \vec{h}_{zp}$

2) Take DFT of X, H

3) Compute $Y[k] = \sqrt{N}\, X[k] \cdot H[k]$

4) Find $\tilde{y}[n] = IDFT\{Y[k]\}$