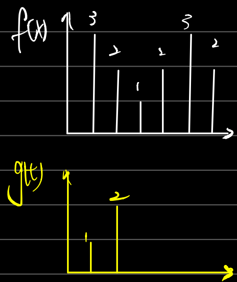
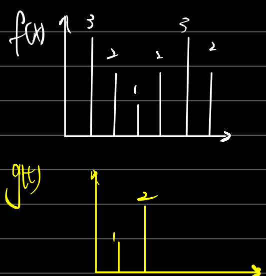
• 2D CNN

Step 1



$1\times3+2\times2=7$

Step 2

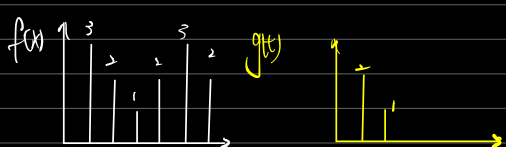


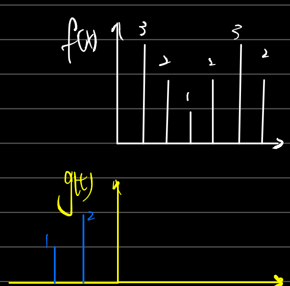
$1\times2+2\times1=5$

• Convolution

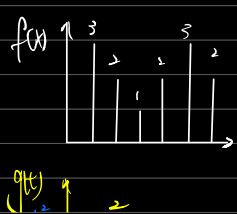
$f(x) * g(x)$

$$(f*g)(w) = \sum_x f(x)\cdot g(w-x)$$
$$= \sum_x f(x)\cdot g(-(x-w))$$

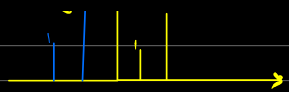




$(f*g)(0)$ (zero-padding)



$(f*g)(2)$



$$(f * g)(w) = \sum_x f(x) \cdot g(w-x) \qquad (f * g)(w) = \int f(x) \cdot g(w-x)\, dx$$

• Convolution Theorem

But it's not clear how to define convolution on graph

$(f * g)(w) = \sum_x f(x) \cdot g(w-x)$ spatial approach

use spectral approach instead.

$$\widehat{(f * g)}(\xi) = \int_{\mathbb{R}} (f * g)(t) \cdot e^{-2\pi i \xi t}\, dt$$

$$= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(u)\, g(t-u)\, du \right] \cdot e^{-2\pi i \xi t}\, dt$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} f(u)\, g(t-u) \cdot e^{-2\pi i \xi t}\, du\, dt$$

$$= \int_{\mathbb{R}} f(u) \left[\int_{\mathbb{R}} g(t-u) \cdot e^{-2\pi i \xi t}\, dt \right] du$$

$$\stackrel{v := t-u}{=} \int_{\mathbb{R}} f(u) \cdot \left[\int_{\mathbb{R}} g(v) \cdot e^{-2\pi i \xi (u+v)}\, dv \right] du$$

$$= \int_{\mathbb{R}} f(u) \cdot \int_{\mathbb{R}} g(v) \cdot e^{-2\pi i \xi u} \cdot e^{-2\pi i \xi v}\, dv\, du.$$

$$= \left(\int_{\mathbb{R}} f(u)\, e^{-2\pi i \xi u}\, du \right) \cdot \left(\int_{\mathbb{R}} f(v)\, e^{-2\pi i \xi v}\, dv \right)$$

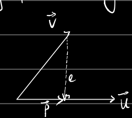
$$= \hat{f}(\xi) \cdot \hat{g}(\xi)$$

$$\widehat{(f * g)}(\xi) = \hat{f}(\xi) \cdot \hat{g}(\xi)$$

Convolution in the spatial domain = multiplication in spectral domain

The dot product $\langle f, g \rangle = \int f(t) \cdot \overline{g(t)}\, dt$

eigen function orthogonality: $\int_0^1 e^{2\pi i m t} \cdot e^{-2\pi i n t}\, dt = \begin{cases} 0 & (m \neq n) \\ 1 & (m = n) \end{cases}$



$$p = \frac{\vec{u}}{\|u\|} \cdot \|\vec{v}\| \cdot \cos\theta$$

$$= \langle u \cdot \vec{v} \rangle \cdot u \quad (\|u\| = |u| = 1)$$

$$\langle f, e^{2\pi i k t} \rangle = \int f(t) \cdot e^{-2\pi i k t}\, dt = \hat{f}(k)$$

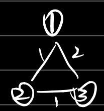
$$f(t) = \sum \hat{f}(k) \cdot e^{2\pi i k t} = \sum \langle f \cdot e^{2\pi i k t} \rangle \cdot e^{2\pi i k t}$$

$(\, p = \langle u, v \rangle \cdot u \,)$

we need a set of orthogonal basis

- Graph Theory.

$G = (V, E, W)$ V: a set of vertices
E: a set of edges
W: weight of edges



$V = \{1, 2, 3\}$
$E = \{(1,2)(2,3)(1,3)\}$
$W = \{1, 1, 2\}$

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

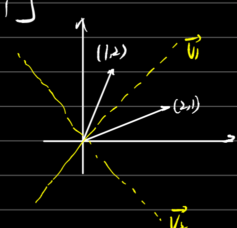
$$D = \mathrm{diag}(\mathrm{sum}(W)) = \begin{bmatrix} 3 & & \\ & 2 & \\ & & 3 \end{bmatrix}$$

$$L_{unnorm} = D - W = \begin{bmatrix} 3 & -1 & -2 \\ -1 & 2 & -1 \\ -2 & -1 & 3 \end{bmatrix} = \begin{bmatrix} W_{12}+W_{13} & -W_{12} & -W_{13} \\ -W_{21} & W_{21}+W_{23} & -W_{23} \\ -W_{31} & -W_{32} & W_{32}+W_{33} \end{bmatrix}$$

$$L_{norm} = D^{-1/2} \cdot L_{unnorm} \cdot D^{-1/2} = L_{unnorm} \cdot \begin{bmatrix} 1/\sqrt{D_1D_1} & 1/\sqrt{D_1D_2} & 1/\sqrt{D_1D_3} \\ 1/\sqrt{D_2D_1} & 1/\sqrt{D_2D_2} & 1/\sqrt{D_2D_3} \\ 1/\sqrt{D_3D_1} & 1/\sqrt{D_3D_2} & 1/\sqrt{D_3D_3} \end{bmatrix}$$

$\because$ L is a $n \times n$ symmetric matrix
$\therefore$ L has n real eigenvalues and n orthogonal eigenvectors.

eg. $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

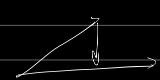


$L = U \Lambda U^T$ where Λ is a diagonal matrix composed of eigenvalues and the column of U is eigenbasis.

$$Lf = U \Lambda U^T f = \begin{bmatrix} | & | & | \end{bmatrix} \begin{bmatrix} \diagdown \end{bmatrix} \begin{bmatrix} - \\ - \\ - \end{bmatrix} \begin{bmatrix} f \end{bmatrix}$$

$(U^T = U^{-1})$

$$= \begin{bmatrix} | & | & | \end{bmatrix} \begin{bmatrix} \diagdown \end{bmatrix} \begin{bmatrix} f \cdot V_1 \\ f \cdot V_2 \\ f \cdot V_3 \end{bmatrix}$$



$f \cdot V_1 = |f| \cdot |V_1| \cdot \cos\theta$
$= |V_1| \cdot |f| \cdot \cos\theta$
$= P_{f, V_1}$

$$= \begin{bmatrix} | & | & | \end{bmatrix} \begin{bmatrix} \lambda_1 P_{V_1} \\ \lambda_2 P_{V_2} \\ \lambda_3 P_{V_3} \end{bmatrix}$$

$$= \lambda_1 P_{V_1} V_1 + \lambda_2 P_{V_2} V_2 + \lambda_3 P_{V_3} V_3$$

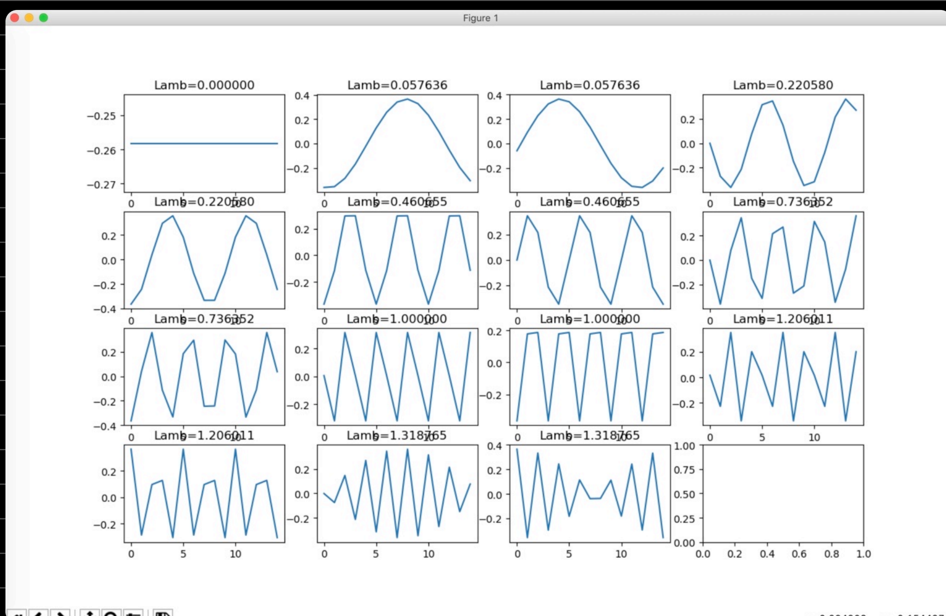
$U^T f$: projection f on eigenvectors of L

$\hat{f}(w) = \int_R f(t) \cdot e^{-2\pi i w t} \, dt$: projection $f(t)$ on eigenfunctions $e^{-2\pi i w t}$

$U U^T f$: reconstruct f from orthogonal components

$f(t) = \int_R \hat{f}(w) \cdot e^{-2\pi i w t} \, dw$: reconstruct f from orthogonal components

eigen value and eigenvector of a circle with 15 nodes 



Define the fourier transform on graph as $\hat{f} = U^T f$
the inverse transform on graph as $f = U \hat{f}$

- Graph convolution

The convolution in spatial domain = multiplication in spectral domain
The spectral decomposition as $U^T f$

convolution $f * g = U(U^T g \odot U^T f)$ where g is the convolution kernel

$X_{l+1} = U(\vec{g} \odot U^T X_l)$

Another way to view kernel g:
g should be a signal in spatial domain, which can be expressed as combination of eigenvectors,

We can view $\vec{g}$ as a function of eigenvalues $\vec{g} = [g(\lambda_0), g(\lambda_1), \dots g(\lambda_{n-1})]$
Thus the convolution becomes

$$X_{l+1} = U(g(\Lambda) \odot U^T X_l)$$

· Chebyshev Polynomial

The graph convolution $X_{l+1} = U(g(\lambda) \odot U^T X_l)$ need to compute the eigenvalue and eigenvectors of L
Instead, we can approximate X_{l+1} using Chebyshev polynomial

1. definition

$$T_n(x) = \cos(n \cdot \arccos(x)) \qquad x := \cos\theta$$

$$\begin{aligned}
T_{n+1}(\cos\theta) = \cos(n\theta + \theta) &= \cos n\theta \cdot \cos\theta - \sin n\theta \cdot \sin\theta \\
&= 2\cos n\theta \cdot \cos\theta - (\cos n\theta \cos\theta + \sin n\theta \sin\theta) \\
&= 2\cos n\theta \cdot \cos\theta - \cos(n\theta - \theta) \\
&= 2 \cdot x \cdot T_n(\cos\theta) - T_{n-1}(\cos\theta)
\end{aligned}$$

$$\therefore \ T_{n+1}(x) = 2x\, T_n(x) - T_{n-1}(x) \quad (x \in [-1,1])$$

2. orthogonality.

$$\int_0^{\pi} \cos mt \cos nt \, dt = \begin{cases} 0 & m \neq n \\ \pi/2 & m = n \neq 0 \\ \pi & m = n = 0 \end{cases}$$

$$\begin{aligned}
&\int_0^{\pi} \cos mt \cos nt \, dt \\
&= \int_1^{-1} \cos(m \cdot \arccos(x)) \cdot \cos(n \cdot \arccos(x)) \, d\arccos x \\
&= \int_{-1}^{1} \frac{T_m(x) \cdot T_n(x)}{\sqrt{1-x^2}} dx
\end{aligned}$$

$\therefore$ $T_k(x)$ gives a set of orthogonal basis in $x \in [-1,1]$
$\therefore$ Any function $f(x)$ can be written as

$$f(x) = \sum_{k=0}^{\infty} A_k T_k(x) = A_0 T_0(x) + A_1 T_1(x) \cdots + A_k T_k(x) \cdots + A_n T_n(x)$$

$$\begin{aligned}
X_{l+1} &= U \cdot g(\Lambda) \cdot U^T X \\
&= U \cdot \left[\sum_{k=0}^{K} \theta_k \cdot T_k(\Lambda)\right] \cdot U^T X \\
&= \sum_{k=0}^{K} \theta_k \cdot U \cdot T_k(\Lambda) \cdot U^T X \\
&= \sum_{k=0}^{K} \theta_k \cdot T_k(L) \cdot X
\end{aligned}$$

Approximate $g(L) X_l$ using Chebyshev polynomial

$$\begin{cases} T_0(x) = 1 \\ T_1(x) = x \\ T_{k+1}(x) = 2x \cdot T_k(x) - T_{k-1}(x) \end{cases}$$

$$\Downarrow$$

$$\{ T_0(L) = I \qquad\qquad \{ T_0(L) X = X$$

$$\begin{cases} T_1(L) = L \\ T_{k+1}(L) = 2 \cdot L \cdot T_k(L) - T_{k-1}(L) \end{cases} \Rightarrow \begin{cases} T_1(L)X = LX \\ T_{k+1}(L)X = 2L \cdot T_k(L)X - T_{k-1}(L)X \end{cases}$$

$$X_{l+1} = \sum_{k=0}^{K} \theta_k T_k(L) \cdot X$$

We can compute the convolution recursively, where θ_k is what we learn by gradient descent

Recap.

① The convolution formula $(f * g)(w) = \sum_x f(x) \cdot g(w-x)$ $\int_x f(x) \cdot g(w-x)\, dx$

② convolution in spatial domain = multiplication in spectral domain $\widehat{(f*g)}(\xi) = \hat{f}(\xi) \cdot \hat{g}(\xi)$

③ generalize the graph Fourier Transform $F(x) = U^T X$ $F^{-1}(x) = UX$

④. graph convolution $X_{l+1} = U \cdot g(\Lambda)\, U^T X_l$

⑤ approximate graph convolution using chebyshev $X_{l+1} = \sum_{k=0}^{K} \theta_k T_k(L) X.$