



中国科学院  
CHINESE ACADEMY OF SCIENCES



# 谱域图卷积介绍

---

报告人：张 奇

中科院自动化研究所

# 目录

---

- **图卷积简介**
- **图谱卷积的背景知识**
  - 谱域图卷积实现思路
  - 拉普拉斯矩阵
  - 图傅里叶变换
  - 卷积定理
- **三个经典图谱卷积模型**
  - SCNN
  - ChebNet
  - GCN

# 图卷积简介

## ➤ 图卷积网络的迅猛发展

➤ 16年以前，图网络相关文献很少。每年仅有1-2篇。

### ICLR 2016

- [Gated Graph Sequence Neural Networks](#)

### ICML 2016

- [Learning Convolutional Neural Networks for Graphs](#)

### CAMD 2016

- [Molecular graph convolutions: moving beyond fingerprints](#)

### NIPS 2015

- [Convolutional Networks on Graphs for Learning Molecular Fingerprints](#)

### ICLR 2014

- [Spectral Networks and Locally Connected Networks on Graphs](#)

### ICCV 2017

- [Situation Recognition with Graph Neural Networks](#)

### ICML 2017

- [Neural Message Passing for Quantum Chemistry](#)
- [Graph-based Isometry Invariant Representation Learning](#)

### IJCAI 2017

- [Knowledge Transfer for Out-of-Knowledge-Base Entities: A Graph Neural Network Approach](#)

### CVPR 2017

- [Geometric deep learning on graphs and manifolds using mixture model CNNs](#)
- [Dynamic Edge-Conditioned Filters in Convolutional Neural Networks on Graphs](#)
- [The More You Know: Using Knowledge Graphs for Image Classification](#)
- [Graph-Structured Representations for Visual Question Answering](#)

### ICLR 2017

- [Semi-Supervised Classification with Graph Convolutional Networks](#)
- [Learning Graphical State Transitions](#)

<https://github.com/naganandy/graph-based-deep-learning-literature/tree/master/conference-publications>

# 图卷积简介

## ➤ 图卷积网络的迅猛发展

## ➤ 2018年，一些会议上大概有了7-8篇图网络文章。

### CIKM 2018 (Oct)

- Semi-supervised Learning on Graphs with Generative Adversarial Nets
- Multiresolution Graph Attention Networks for Relevance Matching
- Heterogeneous Graph Neural Networks for Malicious Account Detection
- Incorporating Corporation Relationship via Graph Convolutional Neural Networks for Stock Price Prediction

### ECCV 2018 (Sep)

- Graph R-CNN for Scene Graph Generation
- Exploring Visual Relationship for Image Captioning
- Learning Human-Object Interactions by Graph Parsing Neural Networks
- Efficient Relative Attribute Learning using Graph Neural Networks
- Local Spectral Graph Convolution for Point Set Feature Learning
- Person Re-identification with Deep Similarity-Guided Graph Neural Network

### UAI 2018 (Aug)

- GaAN: Gated Attention Networks for Learning on Large and Spatiotemporal Graphs

### KDD 2018 (Aug)

- Graph Convolutional Neural Networks for Web-Scale Recommender Systems
- Large-Scale Learnable Graph Convolutional Networks
- DeepInf: Modeling Influence Locality in Large Social Networks
- Adversarial Attacks on Neural Networks for Graph Data
- Deep Variational Network Embedding in Wasserstein Space
- Graph Classification using Structural Attention

### ICML 2018 (Jul)

- Representation Learning on Graphs with Jumping Knowledge Networks
- Stochastic Training of Graph Convolutional Networks
- Learning Steady-States of Iterative Algorithms over Graphs
- GraphRNN: A Deep Generative Model for Graphs
- Adversarial Attack on Graph Structured Data
- NetGAN: Generating Graphs via Random Walks
- Neural Relational Inference for Interacting Systems
- Graph Networks as Learnable Physics Engines for Inference and Control

### IJCAI 2018 (Jul)

- Commonsense Knowledge Aware Conversation Generation with Graph Attention
- Deep Reasoning with Knowledge Graph for Social Relationship Understanding
- Knowledge-Embedded Representation Learning for Fine-Grained Image Recognition
- Adversarially Regularized Graph Autoencoder for Graph Embedding
- Drug Similarity Integration Through Attentive Multi-view Graph Auto-Encoders
- Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting
- Hybrid Approach of Relation Network and Localized Graph Convolutional Filtering for Breast Cancer Subtype Classification

### ACL 2018 (Jul)

- Dating Documents using Graph Convolution Networks
- Semi-supervised User Geolocation via Graph Convolutional Networks
- Graph-to-Sequence Learning using Gated Graph Neural Networks
- A Graph-to-Sequence Model for AMR-to-Text Generation
- Sentence-State LSTM for Text Representation

### CVPR 2018 (Jun)

- Zero-shot Recognition via Semantic Embeddings and Knowledge Graphs
- Multi-Label Zero-Shot Learning with Structured Knowledge Graphs
- Zero-Shot Sketch-Image Hashing
- Image Generation from Scene Graphs
- SplineCNN: Fast Geometric Deep Learning with Continuous B-Spline Kernels
- Dynamic Graph Generation Network: Generating Relational Knowledge from Diagrams
- Efficient Interactive Annotation of Segmentation Datasets with Polygon-RNN++

<https://github.com/naganandy/graph-based-deep-learning-literature/tree/master/conference-publications>

# 图卷积简介

## ➤ 图卷积网络的迅猛发展

➤ 2019年，图网络的文章爆炸性的增长。仅NIPS一个会议就有49篇相关文章。

### NeurIPS 2019 (Dec)

- Understanding the Representation Power of Graph Neural Networks in Learning Graph Topology
- GNN Explainer: A Tool for Post-hoc Explanation of Graph Neural Networks
- Approximation Ratios of Graph Neural Networks for Combinatorial Problems
- Exact Combinatorial Optimization with Graph Convolutional Neural Networks
- HyperGCN: A New Method for Training Graph Convolutional Networks on Hypergraphs
- Powerset Convolutional Neural Networks
- Hyperbolic Graph Convolutional Neural Networks
- [Hyperbolic Graph Neural Networks](#)
- A Flexible Generative Framework for Graph-based Semi-supervised Learning
- Graph-Based Semi-Supervised Learning with Non-ignorable Non-response

▼ more

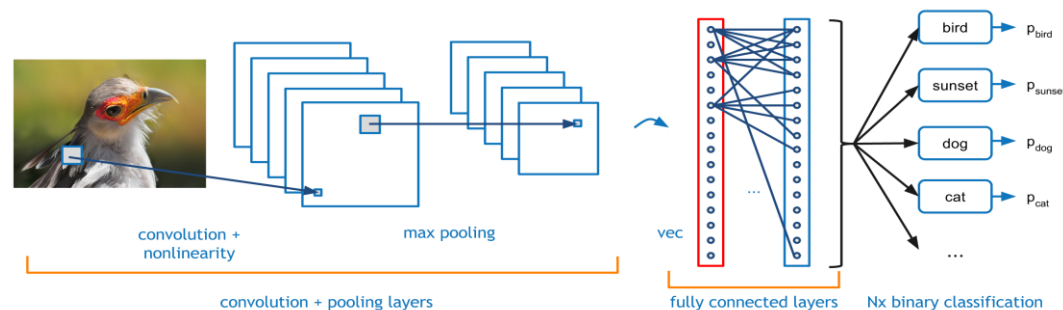
- Graph Agreement Models for Semi-Supervised Learning
- Certifiable Robustness to Graph Perturbations
- On the equivalence between graph isomorphism testing and function approximation with GNNs
- Understanding Attention and Generalization in Graph Neural Networks
- End to end learning and optimization on graphs
- Layer-Dependent Importance Sampling for Training Deep and Large Graph Convolutional Networks
- Graph Transformer Networks
- Diffusion Improves Graph Learning
- Rethinking Kernel Methods for Node Representation Learning on Graphs
- Graph Neural Tangent Kernel: Fusing Graph Neural Networks with Graph Kernels
- Wasserstein Weisfeiler-Lehman Kernels on graphs
- Break the Ceiling: Stronger Multi-scale Deep Graph Convolutional Networks
- Provably Powerful Graph Networks
- Universal Invariant and Equivariant Graph Neural Networks
- DFNetS: Spectral CNNs for Graphs with Feedback-looped Filters

- Stability of Graph Scattering Transforms
- Embedding Symbolic Knowledge into Deep Networks
- D-VAE: A Variational Autoencoder for Directed Acyclic Graphs
- Variational Graph Recurrent Neural Networks
- Efficient Graph Generation with Graph Recurrent Attention Networks
- Conditional Structure Generation through Graph Variational Generative Adversarial Nets
- Graph Normalizing Flows
- Semi-Implicit Graph Variational Auto-Encoders
- N-Gram Graph, A Novel Molecule Representation
- Generative Models for Graph-Based Protein Design
- Heterogeneous Graph Learning for Visual Commonsense Reasoning
- Adaptive GNN for Image Analysis and Editing
- PasteGAN: A Semi-Parametric Method to Generate Image from Scene Graph
- Scene Representation Networks: Continuous 3D-Structure-Aware Neural Scene Representations
- Recurrent Space-time Graph Neural Networks
- Graph-based Discriminators: Sample Complexity and Expressiveness
- Neural Relational Inference with Fast Modular Meta-learning
- Learning Transferable Graph Exploration
- Devign: Effective Vulnerability Identification by Learning Comprehensive Program Semantics via Graph Neural Networks
- Social-BiGAT: Multimodal Trajectory Forecasting using Bicycle-GAN and Graph Attention Networks
- Graph Structured Prediction Energy Networks
- Retrosynthesis Prediction with Conditional Graph Logic Network
- Hyper-Graph-Network Decoders for Block Codes
- Scalable Deep Generative Relational Model with High-Order Node Dependence

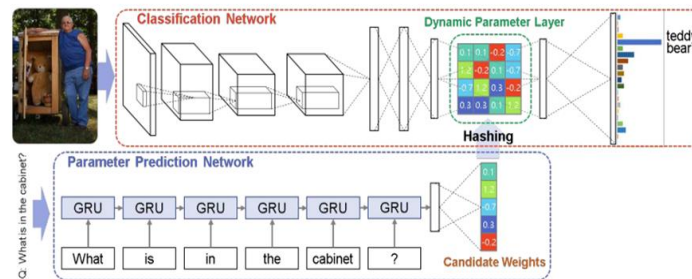
<https://github.com/naganandy/graph-based-deep-learning-literature/tree/master/conference-publications>

# 图卷积简介

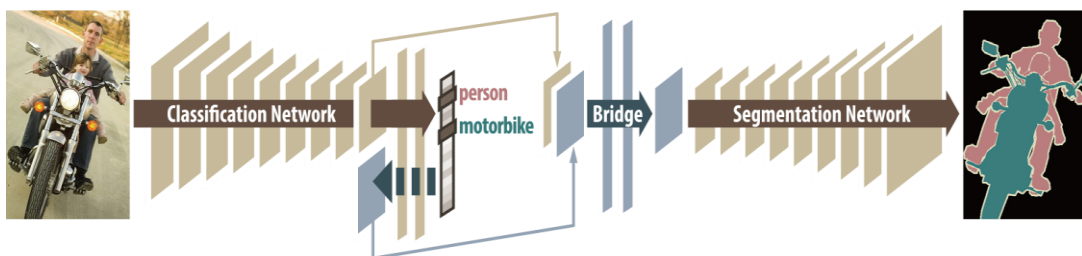
➤ 回顾 经典卷积神经网络已在多个领域取得成功



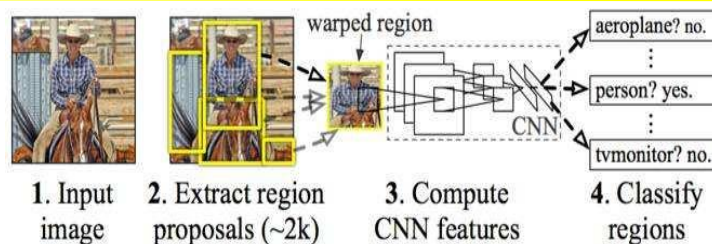
图片  
分类



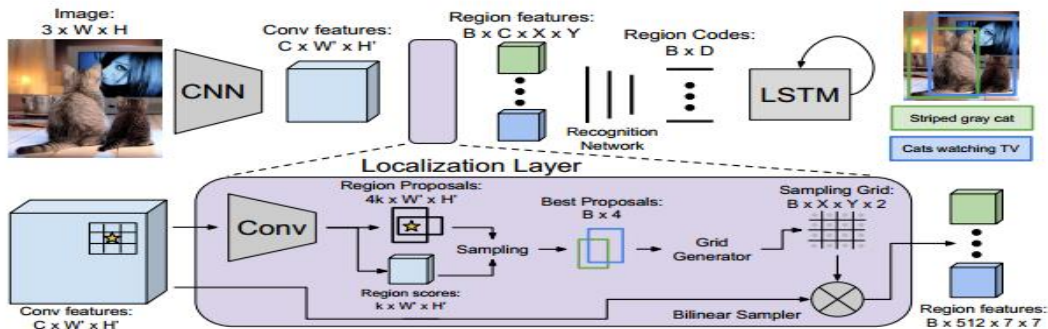
视觉语  
义问答



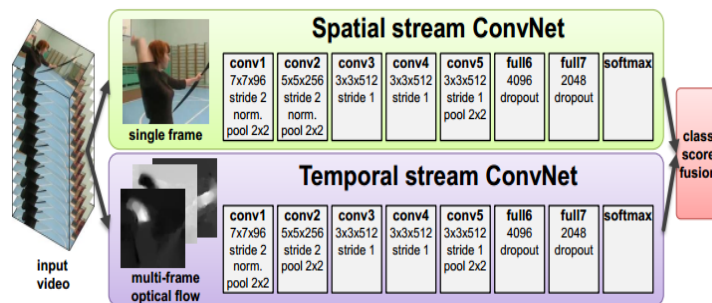
图片  
分割



目标  
检测



看图  
说话

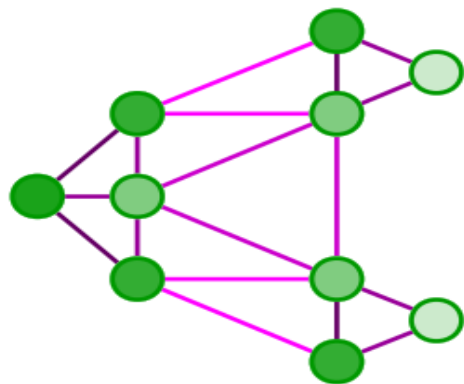
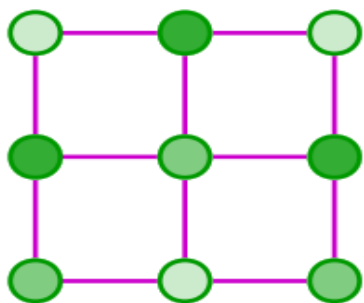


视频  
处理

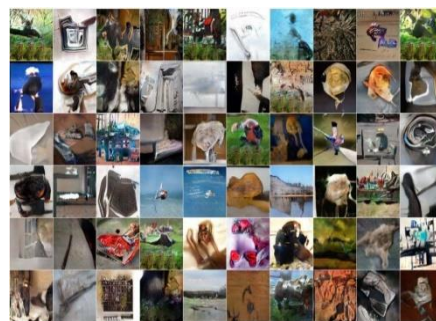


# 图卷积简介

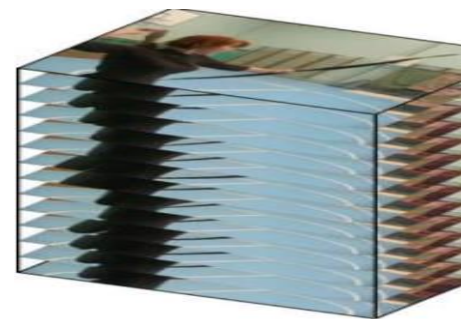
➤ 经典卷积网络的局限：无法处理图结构数据



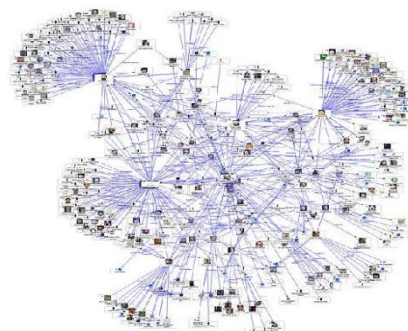
语音



图像



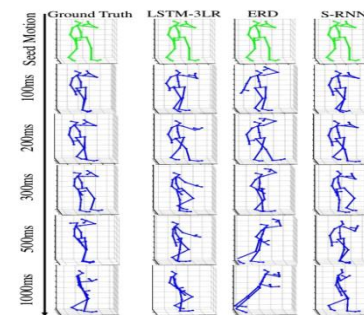
视频



社交网络



分子结构

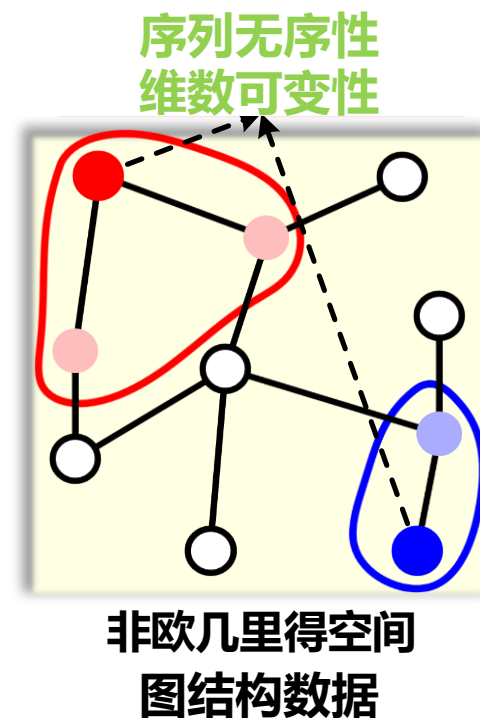
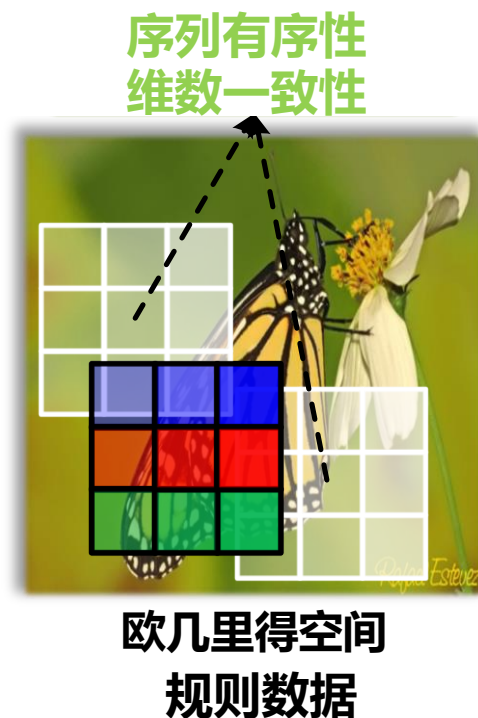


人体骨架

...

# 图卷积简介

- 经典卷积网络的局限：无法处理图结构数据
  - 经典卷积处理图结构数据的局限性
    - 只能处理固定输入维度的数据
    - 局部输入数据必须有序
    - 语音、图像、视频（规则结构）满足以上两点要求。但并不适用于图结构数据（非欧空间数据）



如何将卷积操作扩展到图结构数据中？



# 图卷积简介

## ➤ 图卷积实现思路：

### ➤ 谱域图卷积

- 根据图谱理论和卷积定理，将数据由空域转换到谱域做处理。
- 有较为坚实的理论基础。

本节课重点！

### ➤ 空域图卷积

- 不依靠图谱卷积理论，直接在空间上定义卷积操作。
- 定义直观，灵活性更强。

# 图卷积简介（部分经典模型）

## Spectral approaches 谱域图卷积方法

Kipf & Welling  
(2016)  
GCN

Bruna et al.  
(2014)  
SCNN

Bruna et al.  
(2015)  
SmoothSCNN

Defferrard et al. (2016)  
ChebNet

R.Li et al.  
(2018)  
AGCN

GCNN

Masci et al.  
(2015)

DCNN

Atwood et al.  
(2016)

ACNN

Bocaini et al.  
(2016)

GNN

Hechtlinger  
et al. (2017)

SACNN

Chang et al.  
(2018)

GraphSage

Hamilton et al.  
(2018)

GAT

Velickovic et al.  
(2018)

PGC

Yan et al.  
(2018)

LAGN

Chang et al.  
(2019)

MPNN

Gilmer et al.  
(2017)

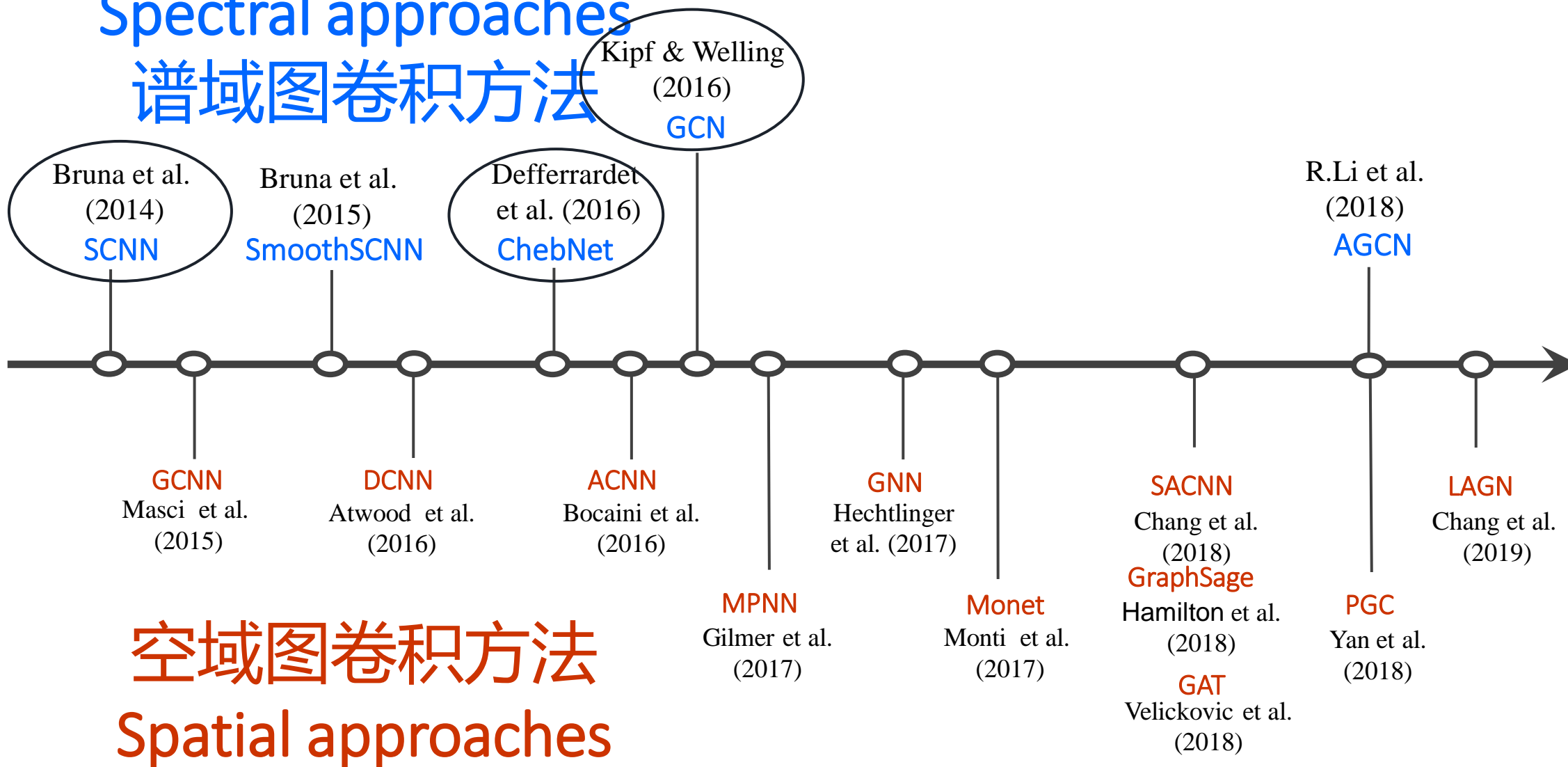
Monet

Monti et al.  
(2017)

## 空域图卷积方法 Spatial approaches

# 图卷积简介（部分经典模型）

## Spectral approaches 谱域图卷积方法



## 空域图卷积方法 Spatial approaches

# 目录

---

- 图卷积简介
- 图谱卷积的背景知识
  - 谱域图卷积实现思路
  - 拉普拉斯矩阵
  - 图傅里叶变换
  - 卷积定理
- 三个经典图谱卷积模型
  - SCNN
  - ChebNet
  - GCN

# 图谱卷积的背景知识

相关知识点：信号处理-卷积定理

## ➤ 1. 谱域图卷积实现思路：

### ➤ 问题1：什么是卷积？

- 根据卷积定理，两信号在空域(或者时域)的卷积的傅里叶变换等于这两信号在频域中的傅里叶变换的乘积，即：

$$\mathcal{F}[f_1(t) \star f_2(t)] = F_1(w) \cdot F_2(w)$$

- $f(t)$  为空域上的信号， $F_1(w)$  为频域上的信号。 $\mathcal{F}$  为傅里叶变换，

$\star$  表示卷积， $\cdot$  为乘积。

- 也可以写成： $f_1(t) \star f_2(t) = \mathcal{F}^{-1}[F_1(w) \cdot F_2(w)]$

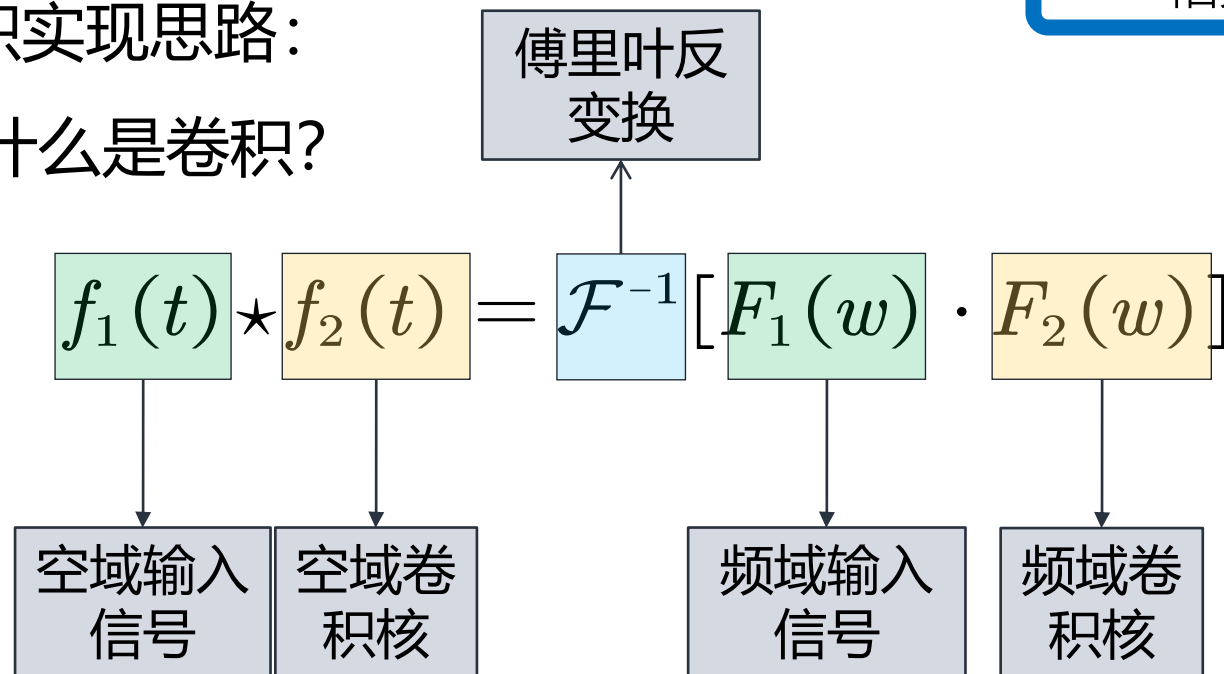
- 其中  $\mathcal{F}^{-1}$  为傅里叶反变换。

# 图谱卷积的背景知识

相关知识点：信号处理-卷积定理

## ➤ 1. 谱域图卷积实现思路：

### ➤ 问题1：什么是卷积？



### ➤ 卷积操作的意义如下：

- 1) 将空域信号转换到频域，然后相乘。
- 2) 将相乘的结果再转换到空域。



# 图谱卷积的背景知识

## ➤ 1. 谱域图卷积实现思路:

### ➤ 问题2: 如何定义图上的傅里叶变换?

基于图谱理论, **图傅里叶变换!**

$$x(t) = \frac{1}{n} \sum_{w=0}^{n-1} e^{i \frac{2\pi}{n} tw} X(w)$$

经典傅里叶变换



$$x(i) = \sum_{l=1}^n \hat{x}(\lambda_l) u_l(i)$$

图傅里叶变换

➤ 我们在接下来的部分会介绍 谱域图卷积的**数学背景知识**, 包括图傅里叶变换。

# 图谱卷积的背景知识

---

## ➤ 2.拉普拉斯矩阵 (Laplacian Matrix)

### ➤ 2.1 拉普拉斯矩阵的定义---符号设置

➤ 无向图  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$   $|\mathcal{V}| = n$

➤ 邻接矩阵  $W \in \mathbb{R}^{n \times n}$

➤ 度矩阵  $D \in \mathbb{R}^{n \times n}$   $D_{ii} = \sum_j W_{ij}$

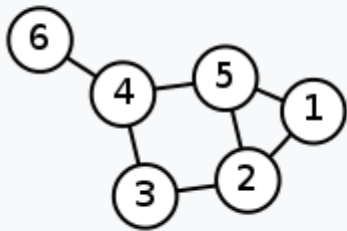
# 图谱卷积的背景知识

相关知识点：图论--拉普拉斯矩阵性质

## ➤ 2.拉普拉斯矩阵

### ➤ 2.1 拉普拉斯矩阵的定义：度矩阵减邻接矩阵

 $G$  $D$  $W$  $L = D - W$ 

Labeled graph	Degree matrix	Adjacency matrix	Laplacian matrix
	$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 2 & -1 & 0 & 0 & -1 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & -1 & 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$

# 图谱卷积的背景知识

相关知识点：  
图论—拉普拉斯矩阵性质

## ➤ 2.拉普拉斯矩阵

### ➤ 2.2 拉普拉斯矩阵是对称半正定矩阵

➤ 半正定证明如下，对任意向量  $f = [f_1, f_2, \dots, f_n]^T$  有

$$\begin{aligned} f^T L f &= f^T D f - f^T W f = \sum_{i=1}^n d_i f_i^2 - \sum_{i,j=1}^n f_i f_j W_{ij} \\ &= \frac{1}{2} \left( \sum_{i=1}^n d_i f_i^2 - 2 \sum_{i,j=1}^n f_i f_j W_{ij} + \sum_{j=1}^n d_j f_j^2 \right) \\ &= \frac{1}{2} \left( \sum_{i,j=1}^n W_{ij} (f_i - f_j)^2 \right) \geq 0 \end{aligned}$$

# 图谱卷积的背景知识

## ➤ 2.拉普拉斯矩阵

### ➤ 2.2 拉普拉斯矩阵是对称半正定矩阵

➤  $L = D - W$   对称半正定矩阵

### ➤ 作为对称半正定矩阵，拉普拉斯矩阵有如下性质：

- n阶对称矩阵一定有n个线性无关的特征向量。
- 对称矩阵的不同特征值对应的特征向量相互正交，这些正交的特征向量构成的矩阵为正交矩阵。
- 实对称矩阵的特征向量一定是实向量
- 半正定矩阵的特征值一定非负。

相关知识点：  
图论—拉普拉斯矩阵性质  
矩阵论—半正定矩阵的性质

# 图谱卷积的背景知识

## ➤ 2.拉普拉斯矩阵

### ➤ 2.3 拉普拉斯矩阵的谱分解

- 特征分解 (Eigen decomposition) , 又称谱分解 (Spectral decomposition)  
是将矩阵分解为由其特征值和特征向量表示的矩阵之积的方法。

相关知识点:

矩阵论—特征分解

矩阵论—半正定矩阵的性质

矩阵论—向量空间

$$L = U \Lambda U^{-1} = U \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} U^{-1}$$

$$U = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$$

$$\vec{u}_i \in \mathbb{R}^n, i = 1, 2, \dots, n$$



# 图谱卷积的背景知识

## ➤ 2.拉普拉斯矩阵

### ➤ 2.3 拉普拉斯矩阵的谱分解

➤  $n$ 阶对称矩阵一定有 $n$ 个线性无关的特征向量（**对称矩阵性质**）。 $n$ 维线性空间中的 $n$ 个线性无关的向量都可以构成它的一组基（**矩阵论知识**）。

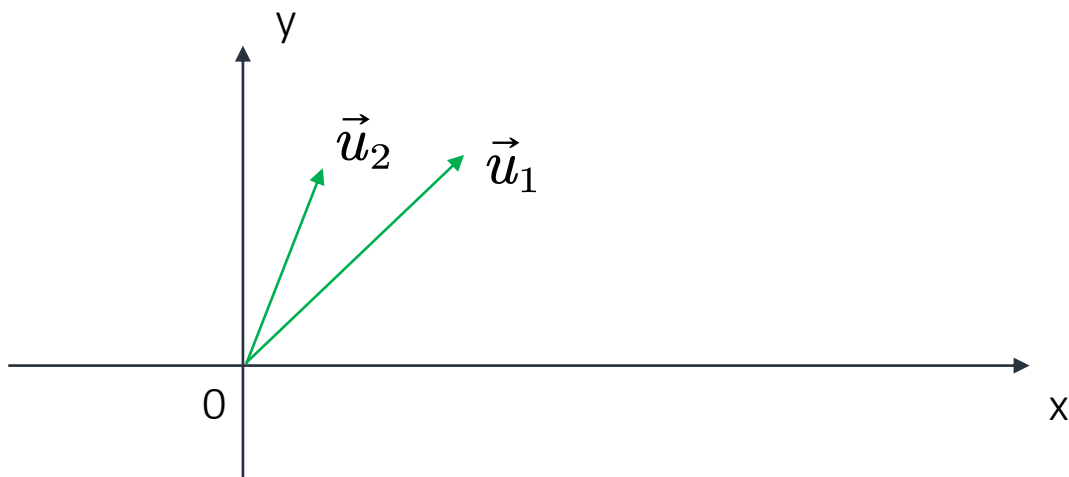
➤ 拉普拉斯矩阵的 $n$ 个特征向量是线性无关的，他们是 $n$ 维空间中的一组基。

相关知识点：

矩阵论—特征分解

矩阵论—半正定矩阵的性质

矩阵论—向量空间



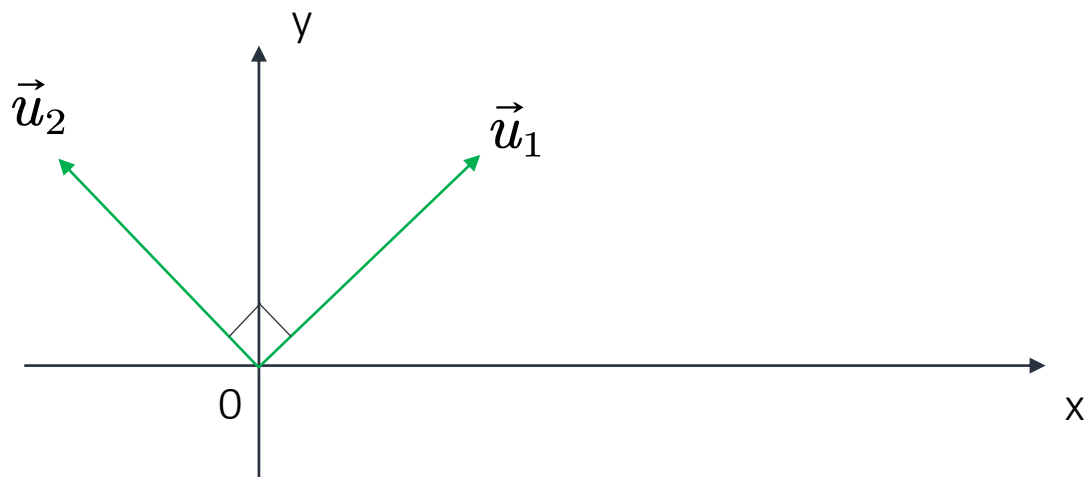
# 图谱卷积的背景知识

## ➤ 2.拉普拉斯矩阵

### ➤ 2.3 拉普拉斯矩阵的谱分解

➤ 对称矩阵的不同特征值对应的特征向量相互正交，这些正交的特征向量构成的矩阵为正交矩阵（对称矩阵性质）

➤ 拉普拉斯矩阵的n个特征向量是n维空间中的一组标准正交基。



相关知识点:

矩阵论—特征分解

矩阵论—半正定矩阵的性质

矩阵论—向量空间

$$UU^T = I$$

$$L = U \Lambda U^{-1} = U \Lambda U^T$$

正交矩阵

# 图谱卷积的背景知识

## ➤ 2.拉普拉斯矩阵

### ➤ 2.4 拉普拉斯矩阵与拉普拉斯算子 (记住结论即可)

➤ 结论：拉普拉斯矩阵是图上的一种拉普拉斯算子。

#### ➤ 2.4.1 拉普拉斯算子 (laplace operator)

➤ 拉普拉斯算子  $\Delta$  的定义为梯度gradient  $\nabla$  的散度divergence  $\nabla \cdot$  。

即：  $\Delta f = \nabla \cdot (\nabla f) = \text{div}(\text{grad}(f))$

➤ 对于n维欧几里得空间，我们可以认为拉普拉斯算子是一个二阶微分算子。

即在各个维度求二阶导数后求和。

$$\Delta f = \sum_i \frac{\partial^2 f}{\partial x_i^2}$$

# 图谱卷积的背景知识

相关知识点：  
数字图像处理

## ➤ 2.拉普拉斯矩阵

### ➤ 2.4 拉普拉斯矩阵与拉普拉斯算子

➤ 在3维欧几里得空间，对于一个三元函数  $f(x, y, z)$  我们可以得到

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

### ➤ 2.4.2 离散情况下欧氏空间的拉普拉斯算子

➤ 离散函数的二阶导数

$$\frac{\partial f}{\partial x} = f'(x) = f(x+1) - f(x)$$

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= f''(x) = f'(x+1) - f'(x) \\ &= f(x+1) + f(x-1) - 2f(x)\end{aligned}$$

# 图谱卷积的背景知识

## ➤ 2.拉普拉斯矩阵

### ➤ 2.4 拉普拉斯矩阵与拉普拉斯算子

- 类似的，对于两个变量的函数  $f(x, y)$  (例如图像)，在y方向也有

$$\frac{\partial^2 f}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y)$$

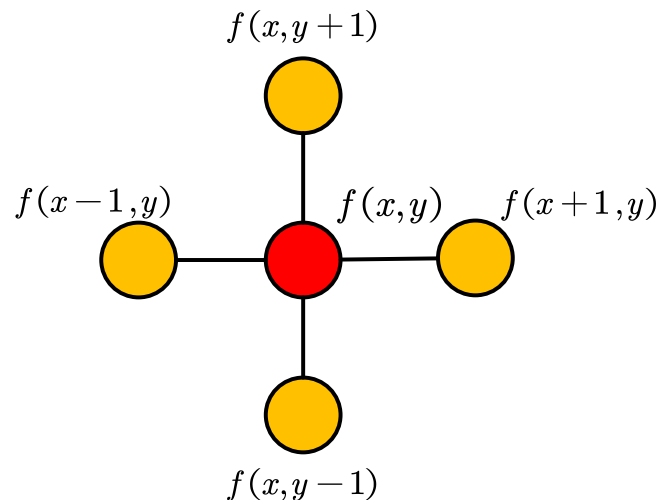
$$\frac{\partial^2 f}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y)$$

- 那么两个变量的离散拉普拉斯算子可以写成

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

$$= f(x+1, y) + f(x-1, y) + f(x, y-1) + f(x, y+1) - 4f(x, y)$$

相关知识点：  
数字图像处理



0	1	0
1	-4	1
0	1	0

冈萨雷斯 数字图像处理

# 图谱卷积的背景知识

## ➤ 2.拉普拉斯矩阵

### ➤ 2.4 拉普拉斯矩阵与拉普拉斯算子

#### ➤ 2.4.3 图上的拉普拉斯算子

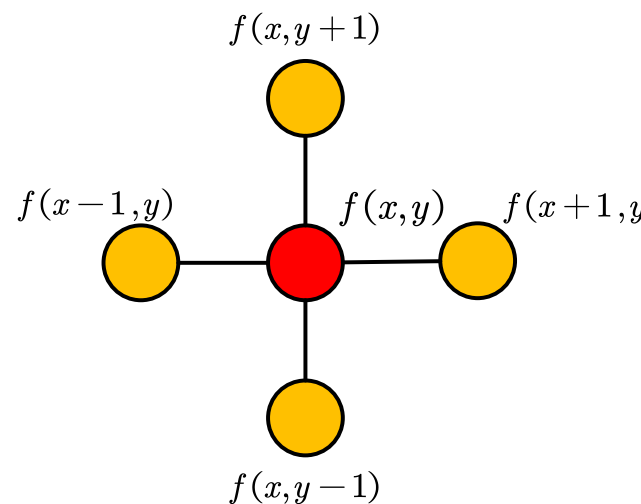
- 欧氏空间内，二维的拉普拉斯算子可以理解为中心节点与周围节点的差值，然后求和

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = f(x+1, y) + f(x-1, y) + f(x, y-1) + f(x, y+1) - 4f(x, y)$$

- 类似的在图上的拉普拉斯算子可以定义如下

$$\Delta f_i = \sum_{(i,j) \in \mathcal{E}} (f_i - f_j)$$

- 其中  $f = (f_1, f_2, \dots, f_n)$ ，代表n个节点上每个节点的信号。





# 图谱卷积的背景知识

## ➤ 2.拉普拉斯矩阵

### ➤ 2.4 拉普拉斯矩阵与拉普拉斯算子

#### ➤ 当有权重时

$$\Delta f_i = \sum_{(i,j) \in \mathcal{E}} W_{ij} (f_i - f_j)$$

➤ 可以理解为中心节点依次减去周围节点，乘以权重后，然后求和。

$$\begin{aligned} \Delta f_i &= \sum_{(i,j) \in \mathcal{E}} W_{ij} (f_i - f_j) \\ &= \sum_{j=1}^n W_{ij} (f_i - f_j) = D_{ii} f_i - \sum_{j=1}^n W_{ij} f_j \end{aligned}$$

#### ➤ 对于n个节点有

$$\begin{aligned} \Delta f = g_\theta &= \begin{pmatrix} \Delta f_1 \\ \vdots \\ \Delta f_n \end{pmatrix} = \begin{pmatrix} D_{11} f_1 - \sum_{j=1}^n W_{1j} f_j \\ \vdots \\ D_{nn} f_n - \sum_{j=1}^n W_{nj} f_j \end{pmatrix} \\ &= \begin{pmatrix} D_{11} & & \\ & \ddots & \\ & & D_{nn} \end{pmatrix} f - Wf = Df - Wf = Lf \end{aligned}$$

# 图谱卷积的背景知识

---

## ➤ 2.拉普拉斯矩阵

### ➤ 2.4 拉普拉斯矩阵与拉普拉斯算子

➤ 2.4.4 结论：拉普拉斯矩阵是图上的一种拉普拉斯算子。

➤ 严谨的证明见 Discrete Regularization on Weighted Graphs for Image and Mesh Filtering 这篇文章。其中定义了在图上如何做梯度运算和散度运算。

➤ 其证明超出本课范围。只要求大家记住这个结论即可。

# 图谱卷积的背景知识

---

## ➤ 2.拉普拉斯矩阵

### ➤ 2.5 拉普拉斯矩阵性质小节

- 拉普拉斯矩阵的 $n$ 个特征向量是 $n$ 维空间中的一组标准正交基
- 实数拉普拉斯矩阵的特征向量一定是实向量。
- 拉普拉斯矩阵的特征值一定非负。
- 拉普拉斯矩阵是图上的一种拉普拉斯算子。

# 图谱卷积的背景知识

相关知识点：信号处理-- 傅里叶变换

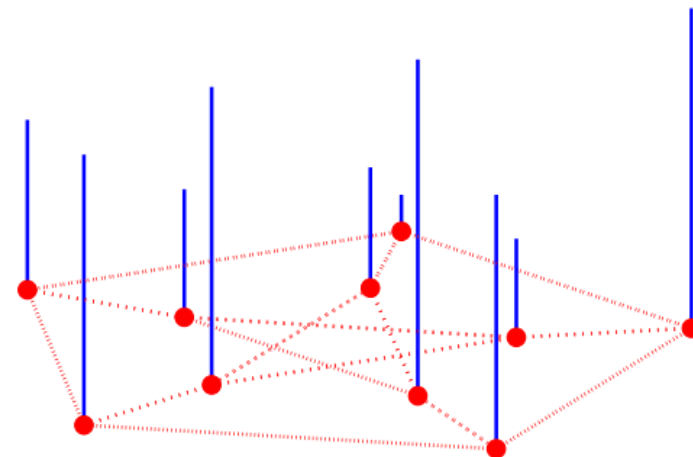
## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 3.1 图上的信号

- 图上的信号一般表达为一个向量。假设有 $n$ 个节点。在这一节，我们将图上的信号记为：

$$x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$$

- 每一个节点上有一个信号值。类似于图像上的像素值。 $i$ 节点上的值为  $x(i) = x_i$



# 图谱卷积的背景知识

相关知识点：信号处理-- 傅里叶变换

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 3.2 经典傅里叶变换

#### ➤ 傅里叶变换的公式：

$$F(w) = \mathcal{F}[f(t)] = \int_{\mathbb{R}} f(t) e^{-2\pi i w t} dt$$

$$f(t) = \mathcal{F}^{-1}[F(w)] = \int_{\mathbb{R}} F(w) e^{2\pi i w t} dw$$

$$F(w) = \sum_{t=1}^n f(t) e^{-i \frac{2\pi}{n} w t}$$
$$f(t) = \frac{1}{n} \sum_{w=1}^n F(w) e^{i \frac{2\pi}{n} w t}$$

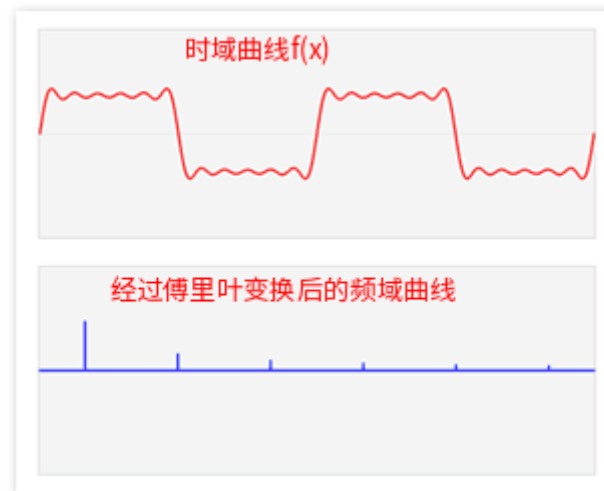
# 图谱卷积的背景知识

相关知识点：信号处理-- 傅里叶变换

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 3.2 经典傅里叶变换

➤ 傅里叶变换中不同频率的余弦函数可视为基函数，其傅里叶系数表示基的振幅。



➤ 相位在这个图中被忽略了，傅里叶系数实际上包含了振幅和相位。



# 图谱卷积的背景知识

相关知识点：信号处理-- 傅里叶变换

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 3.2 经典傅里叶变换

- 傅里叶反变换的本质是：把任意一个函数表示成了若干个正交基函数的线性组合。
- 傅里叶正变换的本质是：求线性组合的系数。具体做法是由原函数和基函数的共轭的内积求得。

$$F(w) = \sum_{t=1}^n f(t) e^{-i \frac{2\pi}{n} wt}$$
$$f(t) = \frac{1}{n} \sum_{w=1}^n F(w) e^{i \frac{2\pi}{n} wt}$$

# 图谱卷积的背景知识

相关知识点：图信号处理

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 3.3 图傅里叶变换

#### ➤ 3.3.1图傅里叶变换定义

- 傅里叶反变换的本质是：把任意一个函数表示成了若干个正交基函数的线性组合。

$$f(t) = \mathcal{F}^{-1}[F(w)] = \int_{\mathbb{R}} F(t) e^{2\pi i w t} dt \quad f(t) = \frac{1}{n} \sum_{w=1}^n F(w) e^{i \frac{2\pi}{n} w t}$$

- 对应graph上的信号  $x \in \mathbb{R}^n$  , 如果要进行一个傅里叶变换, 很自然的我们能想到:  
我们也要找到一组正交基, 通过这组正交基的线性组合来表达  $x \in \mathbb{R}^n$ 。
- 图傅里叶变换, 实际上使用上一小节提到的拉普拉斯矩阵的特征向量, 作为图傅里叶变换的基函数。

$$U = (\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_n)$$

# 图谱卷积的背景知识

相关知识点：图信号处理

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

- 使用拉普拉斯的特征向量作为基函数，任意的图上的信号可以表示为：

$$x = \hat{x}(\lambda_1) \vec{u}_1 + \hat{x}(\lambda_2) \vec{u}_2 + \cdots + \hat{x}(\lambda_n) \vec{u}_n$$

- 为什么任意图上的信号  $x \in \mathbb{R}^n$  都可以表示成这样的线性组合？
  - 原因在于拉普拉斯矩阵的特征向量  $U = (\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_n)$  是n维空间中的n个线性无关的正交向量。
  - n维空间中n个线性无关的向量可以构成空间的一组基（矩阵论知识），而且拉普拉斯矩阵的特征向量还是一组正交基。

# 图谱卷积的背景知识

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 图傅里叶反变换

$$x = \hat{x}(\lambda_1) \vec{u}_1 + \hat{x}(\lambda_2) \vec{u}_2 + \cdots + \hat{x}(\lambda_n) \vec{u}_n$$
$$x(i) = \sum_{l=1}^n \hat{x}(\lambda_l) u_l(i)$$



$$\begin{pmatrix} x(1) \\ x(2) \\ \vdots \\ x(n) \end{pmatrix} = \begin{pmatrix} u_1(1) & u_2(1) & \cdots & u_n(1) \\ u_1(2) & u_2(2) & \cdots & u_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ u_1(n) & u_2(n) & \cdots & u_n(n) \end{pmatrix} \begin{pmatrix} \hat{x}(\lambda_1) \\ \hat{x}(\lambda_2) \\ \vdots \\ \hat{x}(\lambda_n) \end{pmatrix}$$

$$x = U \hat{x}$$

# 图谱卷积的背景知识

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 图傅里叶反变换

$$x = \hat{x}(\lambda_1)\vec{u}_1 + \hat{x}(\lambda_2)\vec{u}_2 + \cdots + \hat{x}(\lambda_n)\vec{u}_n$$
$$x(i) = \sum_{l=1}^n \hat{x}(\lambda_l) u_l(i)$$



$$\begin{pmatrix} x(1) \\ x(2) \\ \vdots \\ x(n) \end{pmatrix} = (\vec{u}_1 \quad \vec{u}_2 \quad \cdots \quad \vec{u}_n) \begin{pmatrix} \hat{x}(\lambda_1) \\ \hat{x}(\lambda_2) \\ \vdots \\ \hat{x}(\lambda_n) \end{pmatrix}$$

$$x = U \hat{x}$$

# 图谱卷积的背景知识

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 图傅里叶正变换

$$\hat{x}(\lambda_l) = \langle x, \vec{u}_l \rangle = \sum_{i=1}^n x(i) u_l(i)$$



$$\begin{pmatrix} \hat{x}(\lambda_1) \\ \hat{x}(\lambda_2) \\ \vdots \\ \hat{x}(\lambda_n) \end{pmatrix} = \begin{pmatrix} u_1(1) & u_1(2) & \cdots & u_1(n) \\ u_2(1) & u_2(2) & \cdots & u_2(n) \\ \vdots & \vdots & \ddots & \vdots \\ u_n(1) & u_n(2) & \cdots & u_n(n) \end{pmatrix} \begin{pmatrix} x(1) \\ x(2) \\ \vdots \\ x(n) \end{pmatrix}$$

$$\hat{x} = U^T x$$

# 图谱卷积的背景知识

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 经典傅里叶变换与图傅里叶变换的对比

$$f(t) = \frac{1}{n} \sum_{w=1}^n F(w) e^{i \frac{2\pi}{n} wt}$$
$$F(w) = \sum_{t=1}^n f(t) e^{-i \frac{2\pi}{n} wt}$$

$$x(i) = \sum_{l=1}^n \hat{x}(\lambda_l) u_l(i)$$
$$\hat{x}(\lambda_l) = \sum_{i=1}^n x(i) u_l(i)$$

### ➤ 为什么基函数一个有共轭，一个没有？

### ➤ 实数拉普拉斯矩阵的特征向量一定是实向量。共轭之后还是自身。

# 图谱卷积的背景知识

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 经典傅里叶变换与图傅里叶变换的对比

#### ➤ 经典傅里叶变换:

$$F(w) = \int_{\mathbb{R}} f(t) e^{-2\pi i w t} dt \text{ (连续)}, \quad X(w) = \sum_{t=1}^n e^{-i \frac{2\pi}{n} t w} x(t) \text{ (离散)}$$

➤ 基:  $e^{2\pi i w t}$  (连续),  $e^{2\pi i w t/n}$  (离散)

➤ 频率:  $w$

➤ 分量的振幅(和相位) :  $F(w)$

#### ➤ 图傅里叶变换:

$$\hat{x}(\lambda_l) = \sum_{i=1}^n x(i) u_l(i)$$

➤ 基:  $\vec{u}_l$

➤ “频率” :  $\lambda_l$

➤ 分量的振幅:  $\hat{x}(\lambda_l)$



# 图谱卷积的背景知识

相关知识点：图信号处理

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 3.3.2 为什么使用拉普拉斯的特征向量作为基 (超出本课范围, 了解即可)

- 经典傅里叶变换有如下规律：傅里叶变换的基函数是拉普拉斯算子的本征函数 (eigenfunctions) (纯数学定理, 掌握结论即可).

$$-\Delta(e^{2\pi i \xi t}) = -\frac{\partial^2}{\partial t^2} e^{2\pi i \xi t} = (2\pi \xi)^2 e^{2\pi i \xi t}.$$

- 又因为拉普拉斯矩阵就是图上的拉普拉斯算子。所以，类似的，图傅里叶变换的基函数即为图拉普拉斯矩阵的特征向量。

$$L\vec{u}_l = \lambda_l \vec{u}_l$$

The Emerging Field of Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and other Irregular Domains.  
IEEE Signal Processing Magazine

# 图谱卷积的背景知识

---

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 3.3.3 特征向量基的性质

- 拉普拉斯矩阵的**特征值**担任了和**频率**类似的位置。
  - 拉普拉斯的特征值都是非负的。且最小特征值为0。类似于经典傅里叶变换中的常数值（可视为频率为0）
- 拉普拉斯矩阵的**特征向量**担任了**基函数**的位置。
  - 0 特征值对应一个常数特征向量，这个和经典傅里叶变换中的常数项类似。
  - 低特征值对应的特征向量比较平滑，高特征值对应的特征向量变换比较剧烈。两者对应于低频基函数和高频基函数。

# 图谱卷积的背景知识

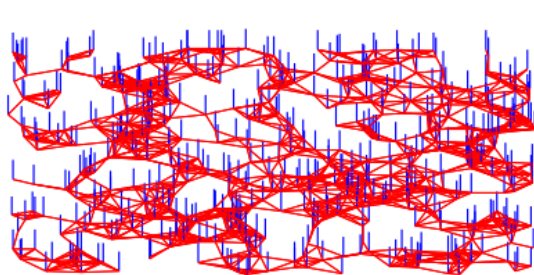
## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 3.3.3 特征向量基的性质

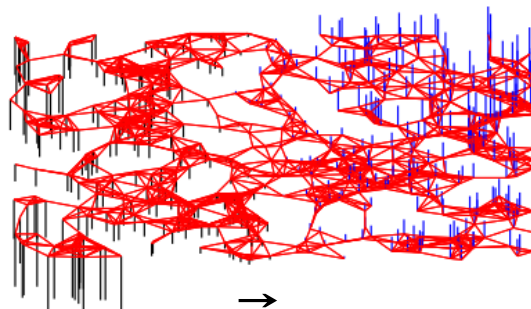
➤ 拉普拉斯矩阵的**特征向量**担任了**基函数**的位置。

➤ 0 特征值对应一个常数特征向量，这个和经典傅里叶变换中的常数项类似。

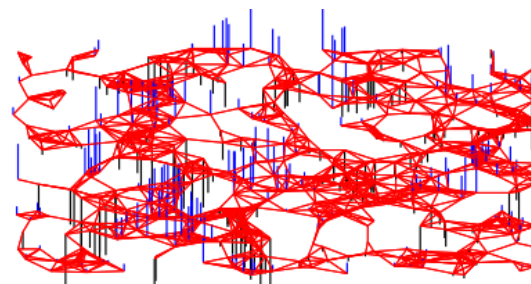
➤ 低特征值对应的特征向量比较平滑，高特征值对应的特征向量变换比较剧烈。两者对应于低频基函数和高频基函数。



$\vec{u}_1$



$\vec{u}_2$



$\vec{u}_{50}$

# 图谱卷积的背景知识

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

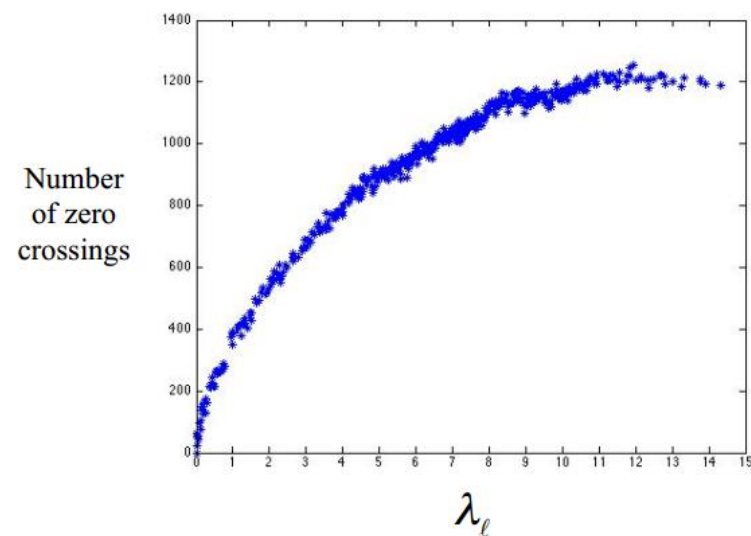
### ➤ 3.3.3 特征向量基的性质

- 通过 “Zero crossing函数” 来的认识特征向量基。

$$\mathcal{Z}_g(\mathbf{x}) = \{e = (i, j) \in \mathcal{E} : x(i)x(j) < 0\}$$

Diagram illustrating the components of the Zero crossing function  $\mathcal{Z}_g(\mathbf{x})$ :

- $\mathbf{x}$  (green box) points to 图结构数据 (Graph structure data) (green box).
- $(i, j) \in \mathcal{E}$  (blue box) points to  $i, j$  节点有边相连 (Nodes  $i, j$  are connected by an edge) (blue box).
- $x(i)x(j) < 0$  (green box) points to 图结构数据在  $i$  节点上的值 (Value of graph structure data at node  $i$ ) (green box).



# 图谱卷积的背景知识

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 3.3.3 特征向量基的性质

- 通过定义图拉普拉斯二次型 (graph Laplacian quadratic form)来定义信号的平滑程度。其表示有边相连的两个节点信号的平方差乘以边的权重后求和。其值越小，代表信号  $x$  越平滑：

$$x^T L x = 1/2 \sum_{i,j=1}^m W_{i,j} (x(i) - x(j))^2$$

相关知识点：瑞利商 定理  
(Rayleigh quotient) 和  
Courant-Fischer Theorem

# 图谱卷积的背景知识

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 3.3.3 特征向量基的性质

$$x^T L x = 1/2 \sum_{i,j=1}^m W_{i,j} (x(i) - x(j))^2$$

$$\vec{u}_l^T L \vec{u}_l = \lambda_l$$

- 当特征向量带入这个函数时, 二次型的值等于特征值。 这恰好是符合经典傅里叶变换中的频率  
设定---频率越高, 基函数 (余弦函数) 变化越陡峭。

# 图谱卷积的背景知识

## ➤ 3.图傅里叶变换 (Graph Fourier Transform )

### ➤ 小结

➤ 依靠图傅里叶变换，可以将定义在graph节点上的信号  $x \in \mathbb{R}^n$  从空间域转换到谱域。

➤ 空间域 --> 谱域  $\hat{x} = U^T x \in \mathbb{R}^n$

➤ 谱域 --> 空间域  $x = U \hat{x}$

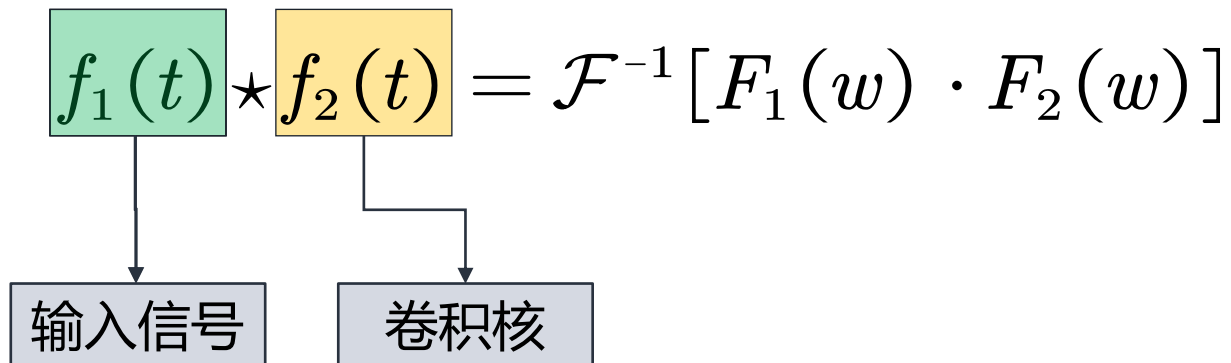
➤ 其中  $L = U \Lambda U^{-1} = U \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} U^{-1}$   $U = (\vec{u}_1, \vec{u}_1, \dots, \vec{u}_n)$

# 图谱卷积的背景知识

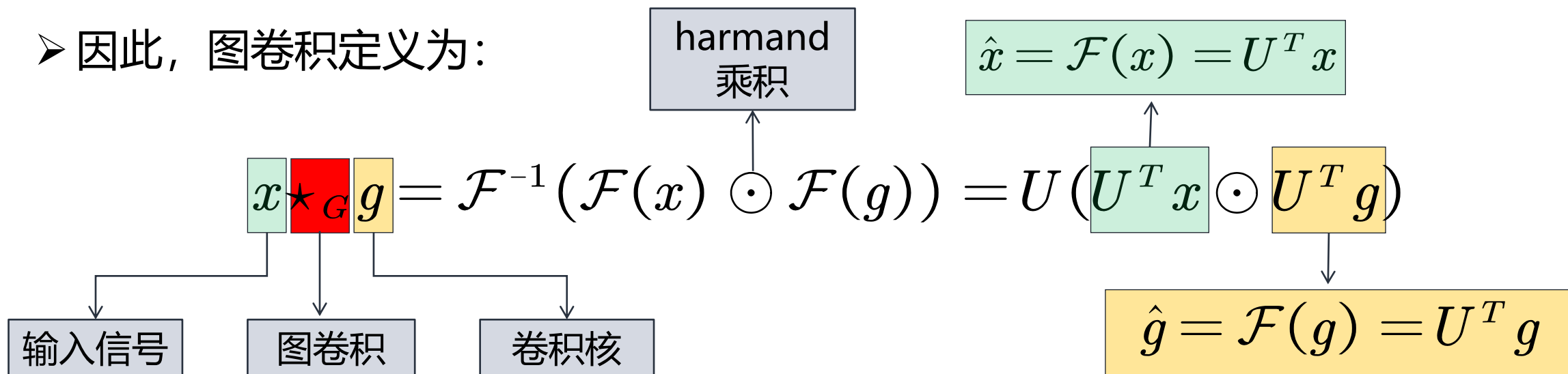
相关知识点：信号处理---卷积定理

## ➤ 4.卷积定理

### ➤ 回顾卷积定理：

$$f_1(t) \star f_2(t) = \mathcal{F}^{-1} [F_1(w) \cdot F_2(w)]$$


### ➤ 因此，图卷积定义为：

$$x \star_G g = \mathcal{F}^{-1} (\mathcal{F}(x) \odot \mathcal{F}(g)) = U (U^T x \odot U^T g)$$


harmand 乘积

$\hat{x} = \mathcal{F}(x) = U^T x$

$\hat{g} = \mathcal{F}(g) = U^T g$

输入信号    图卷积    卷积核



# 图谱卷积的背景知识

相关知识点：信号处理---卷积定理

## ➤ 4.卷积定理

➤ 将两信号分别视为输入信号和卷积核，那么卷积操作可以定义为

- 1) 将空域信号转换到频域，然后相乘。
- 2) 将相乘的结果再转换到空域。

$$\begin{aligned} \begin{pmatrix} \hat{x}(\lambda_1) \\ \hat{x}(\lambda_2) \\ \vdots \\ \hat{x}(\lambda_n) \end{pmatrix} & \quad x \star_G g = \mathcal{F}^{-1}(\mathcal{F}(x) \odot \mathcal{F}(g)) \\ \hat{x} = U^T x & \quad = U (U^T x \odot U^T g) \\ & \quad \downarrow \text{harmand乘积} \\ & \quad \hat{g} = U^T g \end{aligned}$$

# 图谱卷积的背景知识

相关知识点：信号处理---卷积定理

## ➤ 4.卷积定理

➤ 如果以矩阵乘法的形式表达这个公式， 去掉harmand乘积。同时， 通常我们并不关心空间域上的滤波器信号g是什么样子的， 只关心其在频率域的情况。

令  $g_\theta = \text{diag}(U^T g)$ ， 公式等价的转换成下式：

$$x \star_G g_\theta = U g_\theta U^T x = U \begin{pmatrix} \hat{g}(\lambda_1) & & \\ & \ddots & \\ & & \hat{g}(\lambda_n) \end{pmatrix} \begin{pmatrix} \hat{x}(\lambda_1) \\ \hat{x}(\lambda_2) \\ \vdots \\ \hat{x}(\lambda_n) \end{pmatrix} = U \begin{pmatrix} \hat{g}(\lambda_1) & & \\ & \ddots & \\ & & \hat{g}(\lambda_n) \end{pmatrix} U^T x$$

# 目录

---

- 图卷积简介
- 图谱卷积的背景知识
  - 谱域图卷积实现思路
  - 拉普拉斯矩阵
  - 图傅里叶变换
  - 卷积定理
- 三个经典图谱卷积模型
  - SCNN
  - ChebNet
  - GCN

# 三个经典图谱卷积模型

## ➤ 简介

- 三个图谱卷积模型（SCNN、ChebNet、GCN）均立足于谱图理论且一脉相承。
- ChebNet可看做SCNN的改进，GCN可看做ChebNet的改进。
- 三个模型均可认为是下式的一个特例。

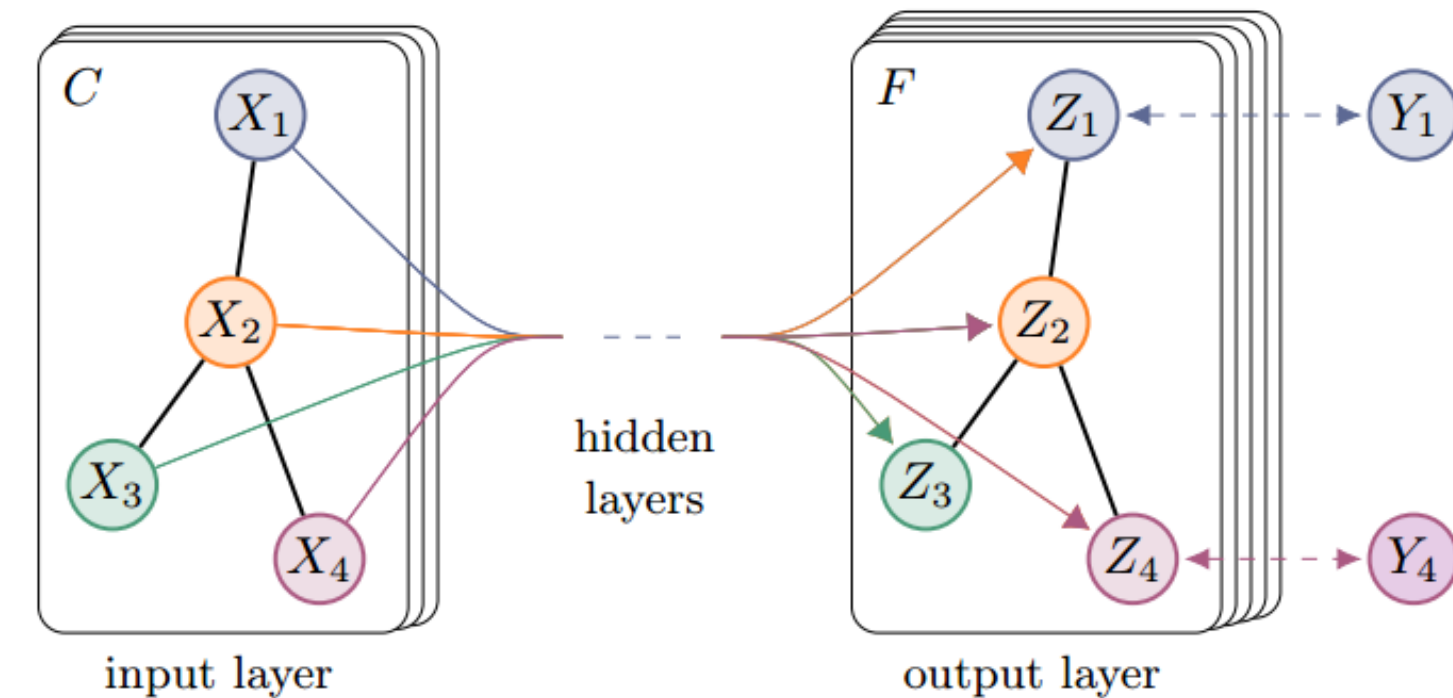
$$x \star_G g_\theta = U g_\theta U^T x = U \begin{pmatrix} \hat{g}(\lambda_1) & & \\ & \ddots & \\ & & \hat{g}(\lambda_n) \end{pmatrix} \begin{pmatrix} \hat{x}(\lambda_1) \\ \hat{x}(\lambda_2) \\ \vdots \\ \hat{x}(\lambda_n) \end{pmatrix}$$

# SCNN

➤ 卷积前后信号的信号:

$$X \in \mathbb{R}^{n \times C}$$

$$Z \in \mathbb{R}^{n \times F}$$



$$X_1 \in \mathbb{R}^{1 \times C}$$

$$Z_1 \in \mathbb{R}^{1 \times F}$$

$$x_{image} \in \mathbb{R}^{h \times w \times C}$$

$h$

1	2	2	5	76	3	8	4	2	9
5	3	36	4	4	23	2	2	66	2
5	1	4	12	42	4	0	1	1	2
6	7	8	3	9	0	1	6	35	1
1	24	5	5	3	23	23	25	20	68
3	12	39	5	32	7	53	9	3	7
1	2	29	3	2	4	23	5	12	6
4	31	7	7	32	3	12	78	88	12
78	68	2	68	8	25	86	12	78	77
35	65	78	12	78	32	57	24	45	14

$w$

# SCNN

## ➤ 核心思想

➤ 用可学习的对角矩阵来代替谱域的卷积核，从而实现图卷积操作。

➤ 即：

$$g_{\theta} = \text{diag}(U^T g)$$



$$g_{\theta} = \text{diag}(\theta)$$



$$g_{\theta} = \begin{pmatrix} \hat{g}(\lambda_1) & & \\ & \ddots & \\ & & \hat{g}(\lambda_n) \end{pmatrix}$$

$$g_{\theta} = \begin{pmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{pmatrix}$$

Spectral networks and locally connected networks on graphs. ICLR, 2014.

# SCNN

## ➤ 卷积定义

➤ 具体公式定义如下：

$$x_{k+1,j} = h \left( U \sum_{i=1}^{C_{k-1}} F_{k,i,j} U^T x_{k,i} \right) \quad (j = 1 \cdots C_k) \quad F_{k,i,j} = \begin{pmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{pmatrix}$$

➤ 其中,  $C_k$  表示第k层的channel (通道) 个数,  $x_{k,i} \in \mathbb{R}^n$  表示第k层的第i个channel的 feature map (特征图) ,

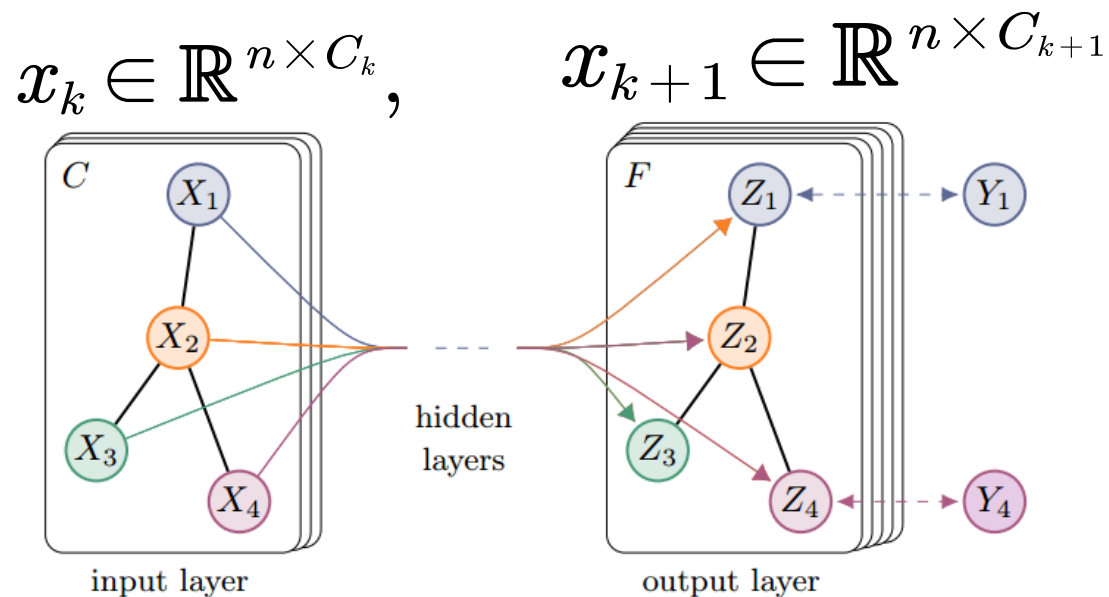
➤  $F_{k,i,j} \in \mathbb{R}^{n \times n}$  代表参数化的谱域的卷积核矩阵。 它是一个对角矩阵, 包含了n个可学习的参数。  $h(\cdot)$  是激活函数。

# SCNN

## ➤ 卷积定义

### ➤ 具体公式定义如下:

$$x_{k+1,j} = h\left(U \sum_{i=1}^{C_k-1} F_{k,i,j} U^T x_{k,i}\right) \quad (j = 1 \cdots C_k)$$



Spectral networks and locally connected networks on graphs. ICLR, 2014.



# SCNN

## ➤ 公式含义

➤ 假设输入 channel 和输出 channel 为 1。简化版本的SCNN卷积公式如下：

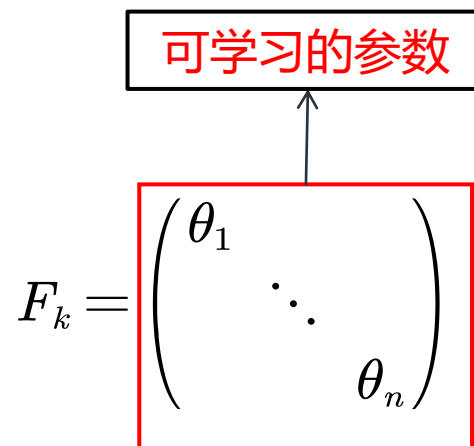
$$x_{k+1} = h(U F_k U^T x_k) \quad x_{k+1} \in \mathbb{R}^n, x_k \in \mathbb{R}^n$$

➤ 这和经典信号处理的频率滤波比较类似：

➤ 空间域 -> 谱域  $U^T x_k$

➤ 对每个“频率分量”放大或者缩小： $F_k U^T x_k$

➤ 谱域 -> 空间域： $U F_k U^T x_k$



The diagram shows a matrix  $F_k$  defined as  $F_k = \begin{pmatrix} \theta_1 & & \\ & \ddots & \\ & & \theta_n \end{pmatrix}$ . The matrix is enclosed in a red box. An arrow points from the top of this box to a black box containing the text '可学习的参数' (Learnable parameters) in red.

# SCNN

---

## ➤ SCNN的缺点

- 计算拉普拉斯矩阵的特征值分解非常耗时。
  - 计算复杂度为  $O(n^3)$ ， $n$ 为节点个数。当处理大规模图数据时（比如社交网络数据，通常有上百万个节点）会面临很大的挑战。
- 模型的参数复杂度较大。
  - 计算复杂度为  $O(n)$ ，当节点数较多时容易过拟合。
- 无法保证局部链接。

# ChebNet

## ➤ 核心思想

- 为了解决SCNN的缺点，ChebNet采用Chebyshev（切比雪夫）多项式代替谱域的卷积核！

$$\boxed{g_\theta = \text{diag}(U^T g)} \xrightarrow{\hat{\Lambda} = \frac{2}{\lambda_{\max}} - I_n} \boxed{g_\theta(\Lambda) = \sum_{k=0}^K \beta_k T_k(\hat{\Lambda})}$$
$$\boxed{g_\theta = \begin{pmatrix} \hat{g}(\lambda_1) & & \\ & \ddots & \\ & & \hat{g}(\lambda_n) \end{pmatrix}} \longrightarrow \boxed{g_\theta = \begin{pmatrix} \sum_{k=1}^K \beta_k T_k(\hat{\lambda}_1) & & \\ & \ddots & \\ & & \sum_{k=1}^K \beta_k T_k(\hat{\lambda}_n) \end{pmatrix}}$$

Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in neural information processing systems (NIPS)*, 2016.

# ChebNet

## ➤ Chebyshev多项式

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) \quad (n = 1, 2, \dots),$$

$$T_0(x) = 1, \quad T_1(x) = x.$$

$$T_2(x) = 2x^2 - 1,$$

$$T_3(x) = 4x^3 - 3x,$$

$$T_4(x) = 8x^4 - 8x^2 + 1,$$

$$T_5(x) = 16x^5 - 20x^3 + 5x,$$

$$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1,$$

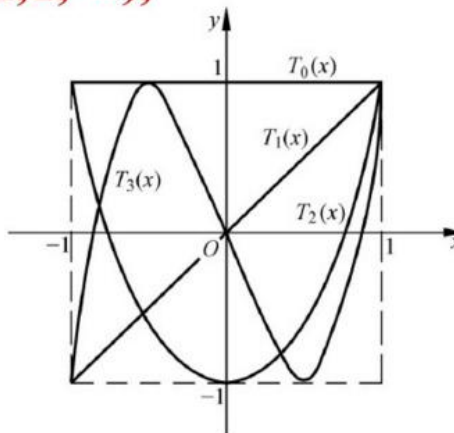


图 3-2

$T_0(x), T_1(x), T_2(x), T_3(x)$  的函数图形见右.

$T_n(x)$  的最高次项系数是  $2^{n-1}$ .

$|T_n(x)| = |\cos n\theta| \leq 1, x \in [-1, 1]$   $x_k = \cos \theta_k = \cos \frac{k}{n}\pi, (k = 0, \dots, n)$  时,

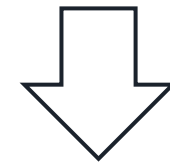
在  $[-1, 1]$  上有  $n+1$  个极值点 有  $|T_n(x_k)| = 1$ ,

$$T_0(L) = I, T_1(L) = L$$

$$T_{n+1}(L) = 2LT_n(L) - T_{n-1}(L)$$

切比雪夫多项式

$$p(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \dots + \beta_K x^K$$



$$p(x) = \beta_0 T_0(x) + \beta_1 T_1(x) + \beta_2 T_2(x) + \dots + \beta_K T_K(x)$$

Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in neural information processing systems (NIPS)*, 2016.

# ChebNet

## ► 理解ChebNet

$$x \star_G g_\theta = U g_\theta U^T x$$

$$= U \sum_{k=0}^K \beta_k T_k(\hat{\Lambda}) U^T x$$

$$= \sum_{k=0}^K \beta_k T_k(U \hat{\Lambda} U^T) x$$

$$= \sum_{k=0}^K \beta_k T_k(\hat{L}) x$$

$$\begin{aligned} & U \sum_{k=0}^K \beta_k T_k(\hat{\Lambda}) U^T \\ &= \sum_{k=0}^K U \beta_k \left( \sum_{c=0}^k \alpha_{kc} \hat{\Lambda}^k \right) U^T \\ &= \sum_{k=0}^K \beta_k \left( \sum_{c=0}^k \alpha_{kc} U \hat{\Lambda}^k U^T \right) \\ &= \sum_{k=0}^K \beta_k \left( \sum_{c=0}^k \alpha_{kc} (U \hat{\Lambda} U^T)^k \right) \\ &= \sum_{k=0}^K \beta_k T_k(U \hat{\Lambda} U^T) \end{aligned}$$

$$\tilde{L} = \frac{2}{\lambda_{\max}} L - I_N$$

Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in neural information processing systems (NIPS)*, 2016.

# ChebNet

## ➤ ChebNet的特点

$$x \star_G g_\theta = U g_\theta U^T x = \sum_{k=0}^K \beta_k T_k(\hat{L}) x$$

Diagram illustrating the equation above. A red box highlights the parameter  $\beta_k$  in the summation. An arrow points from this box to a label above it: **可学习的参数** (Learnable Parameter).

- 1) 卷积核只有K+1个可学习的参数，一般 K远小于n，参数的复杂度被大大降低
- 2) 采用**Chebyshev多项式**代替谱域的卷积核后，经过公示推导，ChebNet不需要对拉普拉斯矩阵做特征分解了。省略了最耗时的步骤。
- 3) 卷积核具有严格的空空间局部性。同时，K就是卷积核的“感受野半径”。即将中心顶点K阶近邻节点作为邻域节点。

Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in neural information processing systems (NIPS)*, 2016.

# GCN

## ➤ 核心思想

- GCN可以视为对ChebNet的进一步简化。仅仅考虑1阶切比雪夫多项式，且每个卷积核仅仅只有一个参数。

由于只考虑1阶切比雪夫多项式，那么只需要两个参数即可！

$$\begin{array}{ccc} \boxed{g_{\theta} = \text{diag}(U^T g)} & \longrightarrow & \boxed{g_{\theta}(\Lambda) = \sum_{k=0}^1 \beta_k T_k(\hat{\Lambda})} \\ \\ \boxed{g_{\theta} = \begin{pmatrix} \hat{g}(\lambda_1) & & \\ & \ddots & \\ & & \hat{g}(\lambda_n) \end{pmatrix}} & \longrightarrow & \boxed{g_{\theta} = \begin{pmatrix} \beta_0 T_0(\hat{\lambda}_1) + \beta_1 T_1(\hat{\lambda}_1) & & \\ & \ddots & \\ & & \beta_0 T_0(\hat{\lambda}_n) + \beta_1 T_1(\hat{\lambda}_n) \end{pmatrix}} \end{array}$$

# GCN

## ➤ 理解GCN

➤ 仅仅考虑一阶切比雪夫多项式。

$$\begin{aligned}x \star_G g_\theta &= \sum_{k=0}^K \beta_k T_k(\hat{L})x = \sum_{k=0}^1 \beta_k T_k(\hat{L})x \\&= \beta_0 T_0(\hat{L})x + \beta_1 T_1(\hat{L})x \\&= (\beta_0 + \beta_1 \hat{L})x \\&= (\beta_0 + \beta_1 (L - I_n))x \\&= (\beta_0 - \beta_1 (D^{-1/2} W D^{-1/2}))x\end{aligned}$$

$$T_0(\hat{L}) = I \quad T_1(\hat{L}) = \hat{L}$$

$$\hat{L} = \frac{2}{\lambda_{max}} L - I_n$$

$$\lambda_{max} = 2$$

$$L = I_n - D^{-1/2} W D^{-1/2}$$



# GCN

## ➤ 理解GCN

➤ 进一步简化, 使得每个卷积核只有一个可学习的参数。

➤ 令  $\beta_0 = -\beta_1 = \theta$

➤ 那么  $x \star_G g_\theta = (\beta_0 - \beta_1 (D^{-1/2} W D^{-1/2})) x = (\theta (D^{-1/2} W D^{-1/2} + I_n)) x$

➤ renormalization trick:

➤ 因为  $D^{-1/2} W D^{-1/2} + I_n$  有范围[0,2]的特征值, 如果在深度神经网络模型中使用该算子, 则反复应用该算子会导致数值不稳定 (发散) 和梯度爆炸/消失, 为了解决该问题, 引入了一个 renormalization trick

# GCN

## ➤ 理解GCN

### ➤ renormalization trick:

$$\begin{aligned} I_n + D^{-1/2} W D^{-1/2} &\rightarrow \tilde{D}^{-1/2} \tilde{W} \tilde{D}^{-1/2} \\ \tilde{W} &= W + I_n \quad \tilde{D}_{ii} = \sum_j \tilde{W}_{ij} \end{aligned}$$

### ➤ 最终公式:

$$x \star_G g_\theta = \theta \left( \tilde{D}^{-1/2} \tilde{W} \tilde{D}^{-1/2} \right) x$$

# GCN

## ➤ GCN的特点

$$x \star_G g_\theta = \theta \left( \tilde{D}^{-1/2} \tilde{W} \tilde{D}^{-1/2} \right) x$$

可学习的参数

- 1) 在忽略input channel 和 output channel的情况下，卷积核只有1个可学习的参数，极大的减少了参数量。（按照作者的说法：“We intuitively expect that such a model can alleviate the problem of overfitting on local neighborhood structures for graphs with very wide node degree distributions, such as social networks, citation networks, knowledge graphs and many other real-world graph datasets.”）
- 2) 虽然卷积核大小减少了（GCN仅仅关注于一阶邻域，类似于3X3的经典卷积），但是作者认为通过多层堆叠GCN，仍然可以起到扩大感受野的作用。

# GCN

---

## ➤ GCN的特点

- 3) 与此同时，**这样极端的参数削减也受到一些人的质疑**。他们认为每个卷积核如果只设置一个可学习参数，会降低模型的能力。（可以参考博文 How powerful are Graph Convolutions? )
  - 如果将传统图像的每一个像素视为graph的一个节点，节点之间为八邻域链接，图像也可以看做一张特殊的图。那么在每个 $3 \times 3$ 的卷积核里，仅仅存在1个可学习的参数。
  - 从目前应用在image的深度学习经验看来，**这样的卷积模型复杂度虽然低，但是模型的能力也遭到了削弱，可能难以处理复杂的任务。**

How powerful are Graph Convolutions? (<https://www.inference.vc/how-powerful-are-graph-convolutions-review-of-kipf-welling-2016-2/>)

# 小结

---

## ➤ 图卷积实现思路：

➤ 现有的图卷积方法大致可以分为 谱域图卷积方法 和 空域图卷积方法。

### ➤ 谱域图卷积

- 根据图谱理论和卷积定理，将数据由空域转换到谱域做处理。
- 有坚实的理论基础。

本节课重点！

### ➤ 空域图卷积

- 不依靠图谱卷积理论。直接在空间上定义卷积操作。
- 定义直观，灵活性更强。

# 小结

---

- 谱域图卷积理论支撑
  - 什么是卷积---卷积定理
  - 如何定义信号的空域/谱域转换---傅里叶变换
  - 图拉普拉斯矩阵的性质---半正定对称矩阵、谱分解
  - 如何定义图上的傅里叶变换---图傅里叶变换

# 小结

## ➤ 三个经典图谱卷积:

### ➤ SCNN

- 用可学习的对角矩阵来代替谱域的卷积核。

$$x \star_G g_\theta = U F U^T x$$

### ➤ ChebNet

- ChebNet采用Chebyshev多项式代替谱域的卷积核。

$$x \star_G g_\theta = U g_\theta U^T x = \sum_{k=0}^K \beta_k T_k(\hat{L}) x$$

### ➤ GCN

- GCN可以视为ChebNet的进一步简化。仅考虑1阶切比雪夫多项式，且每个卷积核仅只有一个参数

$$x \star_G g_\theta = \theta \left( \tilde{D}^{-1/2} \tilde{W} \tilde{D}^{-1/2} \right) x$$

### ➤ 共同特点:

- 均基于卷积定理和图傅里叶变换。

---

谢谢

