Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering

School of Industrial and Management Engineering, Korea University

Jong Kook, Heo

KOREA UNIVERSITY

DMQA

Contents

DMQA

Overview

❖ Introduction

- NeuralIPS 2016 에 게제, 2022년 1월 22일 기준 4405회 인용
- SpectralCNN(Bruna et al, 2014)*을 개량하여 Computational Complexity 를 낮춤
- Localized Spectral Filter 를 제안하였으며, GCN(Kipf et al, 2016)**의 근간이 됨

Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering

Michaël Defferrard Xavier Bresson Pierre Vandergheynst
EPFL, Lausanne, Switzerland
{michael.defferrard,xavier.bresson,pierre.vandergheynst}@epfl.ch

Abstract

In this work, we are interested in generalizing convolutional neural networks (CNNs) from low-dimensional regular grids, where image, video and speech are represented, to high-dimensional irregular domains, such as social networks, brain connectomes or words' embedding, represented by graphs. We present a formulation of CNNs in the context of spectral graph theory, which provides the necessary mathematical background and efficient numerical schemes to design fast localized convolutional filters on graphs. Importantly, the proposed technique offers the same linear computational complexity and constant learning complexity as classical CNNs, while being universal to any graph structure. Experiments on MNIST and 20NEWS demonstrate the ability of this novel deep learning system to learn local, stationary, and compositional features on graphs.

*https://github.com/dudwojae/NeverMind_DMQA/blob/main/GraphNeuralNetworks/20220114/%5B20220114%5DSpectral%20networks%20and%20locally%20connected%20networks%20on%20graphs.pdf

**https://github.com/dudwojae/NeverMind_DMQA/blob/main/GraphNeuralNetworks/20220114/%5B20220114%5DSemi-Supervised%20Classification%20with%20Graph%20Convolutional%20Networks.pdf

Background

❖ Graph Fourier Transform

- Normalized Graph Laplacian Formula

$$L = I_n - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \in \mathbb{R}^{n*n}$$
$$where\ I_n\ is\ the\ identity\ matrix, D\ is\ the\ degree\ matrix$$

- Normalized Graph Laplacian은 **real-symmetric positive semi-definite matrix**
  (1) 특정 행렬이 symmetric 일 경우, eigenvector의 크기는 항상 1이며 서로 직교
  (2) 특정 행렬이 positive semi-definite matrix 일 경우 non-negative eigenvalue 만을 가짐

$$If\ A\ is\ symmetric$$
$$A = V\Lambda V^{-1}\ \&\ A^T = (V\Lambda V^{-1})^T = (V^{-1})^T \Lambda^T V^T$$
$$\therefore V^T = V^{-1} \rightarrow V * V^{-1} = V^T V = V V^T = I$$
$$\therefore v_i \perp v_j\ \forall i, j\ if\ i \neq j,$$
$$|v_i| = 1\ \forall i$$

$$A\ is\ positive\ semi\ definite\ means\ \vec{x}^T A \vec{x} \geq 0\ \forall \vec{x}$$
$$A\vec{v} = \lambda \vec{v}\ by\ definition\ of\ eigenvector$$
$$\therefore \vec{v}^T A \vec{v} = (\vec{v}^T \vec{v})\lambda,$$
$$\vec{v}^T \vec{v}\ is\ necesarily\ positive\ number \rightarrow \lambda\ is\ positive$$

DMQA

Background

❖ Graph Fourier Transform

- Normalized Graph Laplacian Formula

$$L = I_n - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \in \mathbb{R}^{n*n}$$
$$where\ I_n\ is\ the\ identity\ matrix, D\ is\ the\ degree\ matrix$$

- Normalized Graph Laplacian(L) 은 $U\Lambda U^T$ 로 decompose 될 수 있음
  - ✓ $U = [u_0, \dots, u_{n-1}] \in \mathbb{R}^{n*n}$ 은 orthonormal 한 Graph Fourier Basis 의 집합
  - ✓ $\Lambda = \mathrm{diag}([\lambda_0, \dots, \lambda_{n-1}]) \in \mathbb{R}^{n*n}$ 은 입력 신호에 대한 Frequency
  - ✓ **모든 node signal f는 U 의 각 벡터를 basis 로 가지는 frequency domain 으로 매핑 가능**

$$Fourier\ Transform: \ \hat{f} = \{\hat{f}(\lambda_0), \hat{f}(\lambda_1), \dots, \hat{f}(\lambda_{n-1})\} = U^T f$$

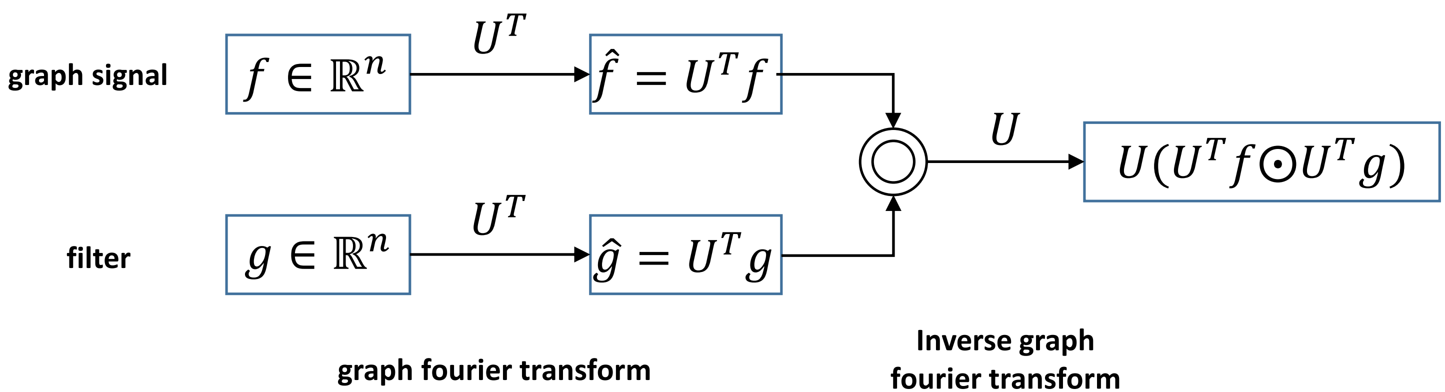
$$Inverse\ Fourier\ Transform: f = U\hat{f} = \Sigma_{i=0}^{n-1} \hat{f}(\lambda_i) u_i$$

DMQA

Background

❖ Graph Fourier Transform

- Spectral Convolution using Graph Fourier Transform
  - ✓ graph signal 과 filter 를 frequency domain 으로 매핑
  - ✓ filter 는 특정 frequency 에 대한 세기를 조절하는 역할
  - ✓ Inverse Fourier Transform 을 통해 original domain 으로 변환



$$Let\ denote\ g_\theta = diag(U^T g) \rightarrow f * g_\theta = U g_\theta U^T f$$

Background

❖ Spectral Convolution

- Spectral Convolution using Graph Fourier Transform
  - ✓ SpectralCNN(Bruna et al, 2014)는 graph filter $g_\theta(\Lambda)$ 를 다음과 같이 정의 하였음

$$y = g_\theta(L)x = g_\theta(U\Lambda U^T)x = \underbrace{\boxed{U g_\theta(\Lambda) U^T}}_{g_\theta(L)} x,$$

$$where\ \boxed{g_\theta(\Lambda) = diag(\theta)}, \theta \in \mathbb{R}^n$$

$$g_\theta(\Lambda) = \begin{bmatrix} g_\theta(\lambda_0) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & g_\theta(\lambda_{n-1}) \end{bmatrix} \Rightarrow g_\theta(\Lambda) = \Theta = \begin{bmatrix} \theta_0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \theta_{n-1} \end{bmatrix}$$

(Bruna et al, 2014)

  - ✓ Spectral CNN Layer 의 Matrix 형태는 다음과 같음

$$H_{:,j}^{(l+1)} = \sigma\left(\Sigma_{i=1}^{f_l} U \Theta_{i,j}^{(l+1)} U^T H_{:,i}^{l}\right), j = (1,2, \dots, f_l)$$

i : input channel index
j : output channel index

DMQA

Proposed Technique

❖ Fast Localized Filter

- Two Limitation of Spectral CNN
  - ✓ They are not localized in space
  - ✓ Learning Complexity is in $O(n)$, the dimensionality of data

- Polynomial Parameterization for localized filter
  - ✓ graph filter $g_\theta(\Lambda)$를 Λ에 대한 **k 차원의 다항 함수**로 정의

$$g_\theta(\Lambda) = \Sigma_{k=0}^{K-1}\theta_k\Lambda^k, \theta \in \mathbb{R}^k$$

$$g_\theta(\Lambda) = \begin{bmatrix} \Sigma_{k=0}^{K-1}\theta_k\lambda_0^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Sigma_{k=0}^{K-1}\theta_k\lambda_{n-1}^k \end{bmatrix}$$

  - ✓ 해당 filter 에 대한 Spectral Convolution 은 아래와 같이 적을 수 있음

$$Ug_\theta(\Lambda)U^T = \Sigma_{k=0}^{K-1}\theta_k U\Lambda^k U^T = \Sigma_{k=0}^{K-1}\theta_k (L)^k = g_\theta(L)$$

  - ✓ 이때, 만약 node i 에서 node j 까지 최단거리가 k-hop 보다 멀다면 다음이 성립함[1],[2]

$$\left[\Sigma_{k=0}^{K-1}\theta_k (L)^k\right]_{i,j} = 0$$

즉, g_θ를 Λ 에 대한 K 차 polynomial로 설정하면
Spatial Domain 에서 K-hop neighbor 만큼만 고려하는 K-localized Filter 가 됨

Proposed Technique

❖ Fast Localized Filter

- Recursive Formulation for fast filtering
  - ✓ Polynomial Parameterization 을 통해 K-Localized Filter 를 만들 수 있음을 확인
  - ✓ **하지만 여전히 연산 복잡도가 $O(n^2)$으로 매우 큼**

- **그래프 필터를 L로부터 재귀적으로 계산되는 다항 함수로 변경**
  - ✓ Chebyshev polynomial $T_k(x)$ 을 이용한 Recursive Formulation
    - ➢ $T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x), with\ T_0 = 1\ and\ T_1 = x$
  - ✓ Diagonal matrix of scaled eigenvalues
    - ➢ $\tilde{\Lambda} = \frac{2\Lambda}{\lambda_{max}} - I_n \rightarrow all\ values\ of\ \tilde{\Lambda}\ lies\ in\ [-1,1],\ \tilde{L} = \frac{2L}{\lambda_{max}} - I_n$
  - ✓ Rescaling 과 Chebyshev polynomial 을 적용한 그래프 필터는 아래와 같이 변경됨

$$g_\theta(\Lambda) = \Sigma_{k=0}^{K-1}\theta_k\Lambda^k, \theta \in \mathbb{R}^k \Rightarrow g_\theta(\Lambda) = \Sigma_{k=0}^{K-1}\theta_k T_k(\tilde{\Lambda})$$

  - ✓ 해당 filter 에 대한 Spectral Convolution 은 아래와 같으며, **연산 복잡도는 $O(K * E)$**

$$Ug_\theta(\Lambda)U^T = U\left(\Sigma_{k=0}^{K-1}\theta_k T_k(\tilde{\Lambda})\right)U^T = \Sigma_{k=0}^{K-1}\theta_k U T_k(\tilde{\Lambda})U^T = \Sigma_{k=0}^{K-1}\theta_k T_k(\tilde{L}) = g_\theta(L)$$

Proposed Technique

❖ Summary

- Spectral Filter 의 변경 과정은 아래와 같이 요약 할 수 있음

$$g_\theta(\Lambda) = g_\theta(\Lambda) = diag(\theta),$$

(Bruna et al, 2014)

$$g_\theta(\Lambda) = \Sigma_{k=0}^{K-1}\theta_k\Lambda^k, \theta \in \mathbb{R}^k$$

Λ에 대한 k차 다항함수로 정의 (K-Localized Filter)

$$g_\theta(\Lambda) = \Sigma_{k=0}^{K-1}\theta_k T_k(\widetilde{\Lambda}), \qquad \widetilde{\Lambda} = \frac{2\Lambda}{\lambda_{max}} - I_n$$

$$T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x), with\ T_0 = 1\ and\ T_1 = x$$

연산 복잡도를 줄이기 위해 ChebyShev 재귀식 도입 (Fast Filter)

DMQA

Conclusion

❖ Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering

- Graph Signal Filter 를 Λ에 대한 k 차 다항 함수로 설정하여 Spatial Domain 에서 Localized Filtering 이 가능함을 보임
- Chebyshev 재귀식을 도입하여 Filter 에 대한 연산복잡도를 줄임
- 해당 논문은 후에 GCN (Kipf et al, 2016) 의 기반이 되는 논문으로써 중요성을 가짐

❖ 소감

- Fast Filtering 을 위해 연산복잡도를 줄이는 것까지는 이해가 되었으나, 수 많은 방법 중 굳이 ChebyShev Recursive Equation 을 쓴 이유는 완벽히 이해하지 못하였음

DMQΛ

Reference

❖ Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering

1. Graph Fourier Transform 설명 : https://ahjeong.tistory.com/15
2. Graph Fourier Transform 설명 : https://sites.icmc.usp.br/gnonato/gs/slide3.pdf