G^2CN: Graph Gaussian Convolution Networks with Concentrated Graph Filters

Mingjie Li, Xiaojun Guo, Yifei Wang,
Yisen Wang, Zhouchen Lin

Spatial Graph Convolution

Let $\boldsymbol{L} = \boldsymbol{I} - \boldsymbol{D}^{-0.5}\boldsymbol{A}\boldsymbol{D}^{-0.5}$ is the graph Laplacian and $\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^\top$ is its eigenvalue decomposition.

Then current Spatial Graph convolutions (like GCN,APPNP and others) can be regarded as graph conovlutions with graph Laplacian's polynomial transformation $g(\boldsymbol{L})$ can be formulated as:

$$\mathbf{H} = g(\boldsymbol{L})\boldsymbol{X} = \boldsymbol{U}g(\boldsymbol{\Lambda})\boldsymbol{U}^\top\boldsymbol{X}$$

with $g(\boldsymbol{\Lambda}) = Diag\big(g(\lambda_1), \dots, g(\lambda_n)\big)$. Since $\widehat{\boldsymbol{X}} = \boldsymbol{U}^\top\boldsymbol{X}$ is graph Fourier Transformation and $\boldsymbol{U}\widehat{\boldsymbol{X}}$ is inverse. We can regard g is the filter function on the graph spectral domain. We call g as graph filters in the following.

However, former analysis on graph filters mainly focus on its global tendency. While ours focus on g's locality properties.

Maximum Response, Centre and Bandwidth

- Maximum Response $\mathcal{R}_g$:

$$\mathcal{R}_g = \max_{\lambda \in [0,2]} |g(\lambda)|$$

- Graph filter g's centre b:

$$g(b) = \mathcal{R}_g$$

- Bandwidth $\mathcal{BW}_g$:

$$\mathcal{BW}_g = \int_0^2 \mathbb{I}\left(g(\lambda), \frac{\mathcal{R}_g}{\sqrt{2}}\right) d\lambda$$

$$\mathbb{I}(\gamma_1, \gamma_2) = \begin{cases} 0, & \gamma_1 < \gamma_2 \\ 1, & \gamma_1 < \gamma_2 \end{cases}$$

Analysis on Different Graph Propagations

Like SGC, we neglect the non-linear activation in our following analysis. The concentration attributes for different models are listed as follows:

Table 8. Concentration Attributes for Bernstein Basis.

Index	0	1	2	3	4	5	6	7	8	9	10
Concentration Center	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0
PassBand	[0.0, 0.069]	[0.079, 0.390]	[0.219, 0.635]	[0.381, 0.855]	[0.555, 1.062]	[0.741, 1.258]	[0.938, 1.445]	[1.145, 1.619]	[1.365, 1.781]	[1.610, 1.921]	[1.931, 2.0]
Maximum Response	1	0.387	0.301	0.267	0.251	0.246	0.251	0.267	0.301	0.387	1

Table 9. Concentration attributes for different graph propagation.

	Graph Propagation Kernel (Or Basis)	$\mathcal{R}$	b	$\mathcal{BW}$
GCN	$(2\mathbf{I}-\mathbf{L})^K$	2^K	0	$2-2^{1-\frac{1}{2K}}$
PPR	$(\mathbf{I}-(1-\alpha)(1-\mathbf{L}))^{-1}$	$\frac{1}{\alpha}$	0	$\frac{(\sqrt{2}-1)\alpha}{1-\alpha}$
ARMA	$\frac{b_K}{1-a_K\mu}$	$\frac{b_K}{1-\lvert a_K\rvert}$	0 or 2	$\frac{(\sqrt{2}-1)(1-\lvert a_K\rvert)}{\lvert a_K\rvert}$
FAGCN	$(1-\lambda+\epsilon)^2$ $(\lambda-1+\epsilon)^2$	$(1+\epsilon)^2$	0 or 2	$(2^{1/4}-1)(1+\epsilon)\in[0.19,0.38]$
Heat Kernel	$e^{-T\mathbf{L}}$	1	0	$\frac{log(\sqrt{2})}{T}$
ChebNet	$\mathbf{C}^{(0)}=\mathbf{I}$ $\mathbf{C}^{(1)}=2\mathbf{L}/\lambda_{\max}-\mathbf{I}$ $\mathbf{C}^{(k)}=2\mathbf{C}^{(2)}\mathbf{C}^{(s-1)}-\mathbf{C}^{(s-2)}$	1	$[0,2]$ $1,-1$	2 $2-\sqrt{2}$ ....
BernNet	$\frac{1}{2^K}\binom{K}{k}(2\mathbf{I}-\mathbf{L})^{K-k}\mathbf{L}^k$	Table 8	$\frac{2}{K}$	Table 8

Weaknesses of Current Graph Propagations:

- **Filter Centres are not flexible:**
  - Filter centres are unchangeable or only have limited choices.
  - Exists in all former models
- **Bandwidth are not flexible:**
  - Bandwidth are unchangeable or only have limited choices.
  - Exists in GCN, ChebNet, BernNet, FAGCN.
- **Bandwidth choice and maximum response are not decoupled:**
  - Exists in PPR, ARMA.

Our Proposed Graph Gaussian Filter

北京大學 PEKING UNIVERSITY

Gaussian Graph Propagation with centre b can be formulated as:

$$\mathbf{H_G} = \mathbf{e}^{-\mathrm{T}(\mathbf{L}-\mathrm{b}\mathbf{I})^2}\mathbf{X}$$

Its graph Filter can be denoted as:

$$g_{G_{T,b}}(\lambda) = e^{-T(\lambda-b)^2}$$

Its centre is b, maximum response is 1 and its bandwidth is:

$$\mathcal{B}W_{g_{G_{T,b}}} = \begin{cases} b, & b < \sqrt{\frac{log\sqrt{2}}{T}} \\ 2\sqrt{\frac{log\sqrt{2}}{T}}, & \sqrt{\frac{log\sqrt{2}}{T}} \le b \le 2 - \sqrt{\frac{log\sqrt{2}}{T}} \\ 2 - b, & b > 2 - \sqrt{\frac{log\sqrt{2}}{T}} \end{cases}$$

Our Proposed Graph Gaussian Convolutions

北京大学 PEKING UNIVERSITY

In practice, we can use K-order Taylor expansion to approximate the Gaussian Graph Filter:

$$\mathbf{H} \approx \sum_{k=0}^{K} \frac{T^k}{k!} (\mathbf{L} - \mathrm{b}\mathbf{I})^{2k} \mathbf{X}$$

Since Gaussian Bases are also a series of universal approximation bases, we use several graph gaussian filters to form our Graph Gaussian Convolution to approximate any graph filter:

$$\mathbf{H} \approx \sum_{i=1}^{N} \theta_i \sum_{k=0}^{K} \frac{T_i^k}{k!} (\mathbf{L} - \mathrm{b}_i\mathbf{I})^{2k} \mathbf{X}$$

That is the formulation of our Graph Gaussian Convolution (G^2Conv) and we can use it to form our Graph Gaussian Convolution Network (G^2CN).

Empirical Results

Table 4. Test accuracy (%) comparison on heterophily datasets. Reported results are averaged over 10 runs.

Dataset	GCN	GAT	APPNP	FAGCN	ARMA	DGC	BernNet	G^2CN
Chameleon	42.98 ± 1.41	47.62 ± 0.83	43.07 ± 1.55	63.3 ± 1.41	55.1 ± 1.32	67.08 ± 1.58	54.11 ± 1.73	**73.61 ± 0.89**
Squirrel	28.41 ± 0.57	27.33 ± 0.81	31.71 ± 0.47	39.7 ± 0.67	35.5 ± 0.81	50.51 ± 0.81	41.35 ± 0.81	**66.91 ± 1.13**
Actor	33.23 ± 1.16	33.93 ± 2.47	39.66 ± 0.55	40.11 ± 0.77	40.79 ± 0.89	40.35 ± 0.73	**41.79 ± 1.01**	41.44 ± 0.76
Texas	77.38 ± 3.28	80.82 ± 2.13	90.98 ± 1.64	96.5 ± 0.47	92.3 ± 0.66	94.74 ± 0.33	96.22 ± 0.79	**96.72 ± 0.73**
Cornell	65.90 ± 4.43	78.21 ± 2.95	91.81 ± 1.96	93.3 ± 1.21	93.4 ± 1.13	92.45 ± 1.21	92.29 ± 2.74	**94.11 ± 1.81**

You can obtain other empirical results and analysis in our paper.

Thanks for Watching！