Summary Questions of the last lecture

- Explain the key idea of Spectral GCN: Spectral Networks and Deep Locally Connected Networks on Graphs (Bruna et al. ICLR 2014)

→ For spectral filtering of single channel graph signal, we need N spectral filters which depend on eigenvalues of graph Laplacian. The spectral GCN parameterizes each spectral filter for an eigenvalue by single learnable parameter. Thus the number of total learnable parameters become N for single channel graph signal and $d_1 \times d_2 \times N$ for multi-channel graph signal with d_1 input channels and d_2 output channels.

$\Longrightarrow d_1 \times d_2 \times K$ parameters

$$\hat{g}_{nm}(\Lambda) = \begin{bmatrix} \hat{g}_{nm}(\lambda_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \hat{g}_{nm}(\lambda_N) \end{bmatrix}$$

$$\hat{g}_{nm}(\Lambda) = \begin{bmatrix} \theta_1^{(nm)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \theta_N^{(nm)} \end{bmatrix} \Longrightarrow \hat{g}_{nm}(\boldsymbol{L}) = \boldsymbol{U}\hat{g}_{nm}(\Lambda)\boldsymbol{U}^{\top}$$

Summary Questions of the last lecture

- Explain the key idea of ChebNet: Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering (Defferard et al. NIPS 2016)

→ The ChebNet: parameterizes each spectral filter by a high order polynomial of an eigenvalue where the coefficients become learnable parameters and the parameters are shared for all eigenvalues. The high order spectral filters are provable to be strictly localized in a ball of radius K, i.e. K hops from the central vertex. Note $K << N$.

$\Longrightarrow d_1 \times d_2 \times K$ parameters

$$\boldsymbol{F}_O(:,n) = \boldsymbol{U}\hat{g}_{nm}(\Lambda)\boldsymbol{U}^T\boldsymbol{F}_I(:,m) = \sum_{m=1}^{d_1}\sum_{k=0}^{K-1}\theta_k^{(nm)}\,\boldsymbol{L}^k\,\boldsymbol{F}_I(:,m)$$

$$\hat{g}_{nm}(\Lambda) = \begin{bmatrix} \sum_{k=0}^{K-1}\theta_k^{(nm)}\lambda_1^k & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sum_{k=0}^{K-1}\theta_k^{(nm)}\lambda_N^k \end{bmatrix} \Longrightarrow \hat{g}_{nm}(\boldsymbol{L}) = \boldsymbol{U}\hat{g}_{nm}(\Lambda)\boldsymbol{U}^T$$

Summary Questions of the last lecture

- Explain the key idea of ChebNet: Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering (Defferard et al. NIPS 2016)

→ The ChebNet parameterizes each spectral filter by Chebyshev polynomial of an eigenvalue where the coefficients become learnable parameters and the parameters are shared for all eigenvalues. Thus the number of total learnable parameters become the number of coefficients, K, for single channel graph signal and $d_1 \times d_2 \times K$ for multi-channel graph signal with d_1 input channels and d_2 output channels. Chebyshev polynomials form an orthogonal basis for the Hilbert space.

Multiple channel: $d_2 \times d_1 \times K$ parameters

$$\boldsymbol{F}_O(:,n) = \sum_{m=1}^{d_1} \sum_{k=0}^{K} \theta_k^{(mn)} T_k(\tilde{\boldsymbol{L}}) \boldsymbol{F}_I(:,m) \qquad \tilde{\boldsymbol{L}} = \frac{2L}{\lambda_{max}} - \mathbf{I}$$

Summary Questions of the last lecture

- Explain the key idea of Simplified ChebNet: Semi-Supervised Classification with Graph Convolutional Networks (Kipf & Welling, ICLR 2017)

→ The Simplified ChebNet parameterizes each spectral filter by the first order Chebyshev polynomial of an eigenvalue where single coefficient becomes a learnable parameter and the parameter is shared for all eigenvalues. The Chebyshev filter matrix is renormalized to have eigenvalues in [0,1].

Single channel: $m = 1, n = 1$

$$\boldsymbol{f}_O = \theta\left(\widetilde{\boldsymbol{D}}^{-1/2}\widetilde{\boldsymbol{A}}\widetilde{\boldsymbol{D}}^{-1/2}\right)\boldsymbol{f}_I, \text{ with } \widetilde{\boldsymbol{A}} = \boldsymbol{A} + \mathbf{I}, \text{ and } \widetilde{\boldsymbol{D}}_{ii} = \sum_j \widetilde{\boldsymbol{A}}_{ij}$$

Multiple channel: $m = 1, \dots, d_1; \;\; n = 1, \dots, d_2$

$$\boldsymbol{F}_O(:, n) = \sum_{m=1}^{d_1} \theta^{(mn)}\left(\widetilde{\boldsymbol{D}}^{-1/2}\widetilde{\boldsymbol{A}}\widetilde{\boldsymbol{D}}^{-1/2}\right)\boldsymbol{F}_I(:, m)$$