LEARNING GRAPH STRUCTURE FROM CONVOLUTIONAL MIXTURES

A NOTE

Zepeng Zhang

June 18, 2022

This paper [1] proposed the graph deconvolution network (GDN), a neural network unrolled from proximal gradient descent that is capable of recovering latent graph structure from observations of its convolutional mixtures, i.e., related graphs containing spurious, indirect relationships.

The problem is to recover the sparse latent graph $\mathbf{A}_L$ from the observed graph $\mathbf{A}_O = \sum_{k=0}^{K} h_k \mathbf{A}_L^k$ by solving

$$\hat{\boldsymbol{A}}_L \in \underset{\boldsymbol{A}\in\mathcal{A}}{\arg\min} \left\{ \|\boldsymbol{A}\|_1 + \frac{\lambda}{2} \|\boldsymbol{A}_O - \boldsymbol{H}(\boldsymbol{A};\boldsymbol{h})\|_F^2 \right\},$$

where $\mathcal{A} = \left\{ \mathbf{A} \in \mathbb{R}^{N\times N} \mid \operatorname{diag}(\mathbf{A}) = \mathbf{0}, A_{ij} = A_{ji} \geq 0, \forall i,j, \in \{1,\dots,N\} \right\}$.

Denoting $g(\mathbf{A}) = \|\boldsymbol{A}_O - \boldsymbol{H}(\boldsymbol{A};\boldsymbol{h})\|_F^2$, the proximal gradient iterations can be expressed as

$$\boldsymbol{A}[k+1] = \operatorname{ReLU}\left(\boldsymbol{A}[k] - \tau \nabla g(\boldsymbol{A}[k]) - \tau \mathbf{1}\mathbf{1}^\top\right).$$

By approximating the gradient with its first-order approximation

$$\begin{aligned}\nabla g(\boldsymbol{A}) &= -\sum_{k=1}^{K} h_k \sum_{r=0}^{k-1} \boldsymbol{A}^{k-r-1} \boldsymbol{A}_O \boldsymbol{A}^r + \frac{1}{2} \nabla_{\boldsymbol{A}} \operatorname{Tr}\left[\boldsymbol{H}^2(\boldsymbol{A};\boldsymbol{h})\right] \\ &\approx -h_1 \boldsymbol{A}_O - h_2 \left(\boldsymbol{A}_O \boldsymbol{A} + \boldsymbol{A}\boldsymbol{A}_O\right) + \left(2h_0 h_2 + h_1^2\right) \boldsymbol{A},\end{aligned}$$

we have

$$\boldsymbol{A}[k+1] = \operatorname{ReLU}\left(\alpha \boldsymbol{A}[k] + \beta \left(\boldsymbol{A}_O \boldsymbol{A}[k] + \boldsymbol{A}[k]\boldsymbol{A}_O\right) + \gamma \boldsymbol{A}_O - \tau \mathbf{1}\mathbf{1}^\top\right).$$

The unknown filter coefficients are then treated as learnable parameters $\boldsymbol{\Theta} := \{\alpha, \beta, \gamma, \tau\}$, which can be shared or independent across layers. A MIMO architecture can also be used, in which the input and output can be tensors and multiple sets of parameters are learned.

References

[1] Max Wasserman, Saurabh Sihag, Gonzalo Mateos, and Alejandro Ribeiro. Learning graph structure from convolutional mixtures. *arXiv preprint arXiv:2205.09575*, 2022.