Notes

Mingyu Liu

2021/12/01

1. Deep latent Gaussian models (DLGMs)

A general class of deep directed graphical models that consist of Gaussian latent variables at each layer of a processing hierarchy. Descend through the hierarchy and generate observations $\mathbf{v}$ by sampling from the observation likelihood using the activation of the lowest layer $\mathbf{h}_1$.

$$\begin{aligned}
\boldsymbol{\xi}_l &\sim \mathcal{N}\left(\boldsymbol{\xi}_l \mid \mathbf{0}, \mathbf{I}\right), \quad l = 1, \ldots, L \\
\mathbf{h}_L &= \mathbf{G}_L \boldsymbol{\xi}_L \\
\mathbf{h}_l &= T_l\left(\mathbf{h}_{l+1}\right) + \mathbf{G}_l \boldsymbol{\xi}_l, \quad l = 1 \ldots L-1 \\
\mathbf{v} &\sim \pi\left(\mathbf{v} \mid T_0\left(\mathbf{h}_1\right)\right),
\end{aligned}$$

where $\boldsymbol{\xi}_l$ are mutually independent Gaussian variables. The transformations T_l represent multi-layer perceptrons (MLPs) and $\mathbf{G}_l$ are matrices. At the visible layer, the data is generated from any appropriate distribution $\pi(\mathbf{v} \mid \cdot)$ whose parameters are specified by a transformation of the first latent layer. The maps T_l and the matrices G_l are parametrized by $\boldsymbol{\theta}^g$ (weak Gaussian prior over $p\left(\boldsymbol{\theta}^g\right) = \mathcal{N}(\boldsymbol{\theta} \mid \mathbf{0}, \kappa \mathbf{I})$).