

Repulsive Variational Gaussian Mixture Models

Omid S. Solari*

SOLARI@BERKELEY.EDU

*Department of Statistics
University of California, Berkeley*

Yulun Wu*

YULUN_WU@BERKELEY.EDU

*Department of Statistics
University of California, Berkeley*

James B. Brown

JBBROWN@LBL.GOV

*Lawrence Berkeley National Laboratory and Department of Statistics
University of California, Berkeley*

Editor:

Abstract

tentative achievements:

1. True Gaussian mixture model
2. Bayesian repulsive priors
3. cluster single cell data and map them to body

Keywords: Gaussian Mixtures, Clustering, Variational Auto-Encoders

1. Introduction

- review use cases of gmms
- review uses in genomics specifically

2. Background

- review gaussian mixture models
- review vae
- introduce repulsive priors
- set up the use for single cell data

mention the shortcomings of GMVAE.

*. equally contributing authors, alphabetical order

3. Problem Statment

general definition of Gaussian mixture models

4. Model

4.1 Gaussian Mixture Variational Auto-Encoder

Dependency structure: (draw a graph)

4.1.1 VARIATIONAL LOWER BOUND

$$\mathcal{L}_{ELBO} = \mathbb{E}_{q_\phi(\mathbf{v}_{latent}|\mathbf{v}_{observed})} \left[\log \frac{p_\theta(\mathbf{v}_{latent}, \mathbf{v}_{observed})}{q_\phi(\mathbf{v}_{latent}|\mathbf{v}_{observed})} \right] \quad (1)$$

4.1.2 SUPERVISED

TODO: define notations

$$\begin{aligned} \mathcal{D} [q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y}) \parallel p_\theta(\mathbf{z}|\mathbf{x}, \mathbf{y})] &= -\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y})} \left[\log \frac{p_\theta(\mathbf{x}, \mathbf{y}, \mathbf{z})}{q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y})} \right] + \log p_\theta(\mathbf{x}, \mathbf{y}) \\ &= -\mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y})} [\log p_\theta(\mathbf{x}|\mathbf{z}) + \log p_\theta(\mathbf{z}|\mathbf{y}) - \log q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y})] \\ &\quad - \log \frac{p(\mathbf{y})}{q_\phi(\mathbf{y}|\mathbf{x})} - \log q_\phi(\mathbf{y}|\mathbf{x}) + \log p_\theta(\mathbf{x}, \mathbf{y}) \end{aligned} \quad (2)$$

$$\begin{aligned} \mathcal{L}_{ELBO} &= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y})} [\log p_\theta(\mathbf{x}|\mathbf{z})] + \log q_\phi(\mathbf{y}|\mathbf{x}) \\ &\quad - \mathcal{D} [q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y}) \parallel p_\theta(\mathbf{z}|\mathbf{y})] - [q_\phi(\mathbf{y}|\mathbf{x}) \parallel p(\mathbf{y})] \end{aligned} \quad (3)$$

(Sometimes people add term $\alpha \log q_\phi(\mathbf{y}|\mathbf{x})$ to amplify classification loss.)

4.1.3 UNSUPERVISED

$$\mathcal{D} [q_\phi(\mathbf{y}, \mathbf{z}|\mathbf{x}) \parallel p_\theta(\mathbf{y}, \mathbf{z}|\mathbf{x})] = -\mathbb{E}_{q_\phi(\mathbf{y}, \mathbf{z}|\mathbf{x})} \left[\log \frac{p_\theta(\mathbf{x}, \mathbf{y}, \mathbf{z})}{q_\phi(\mathbf{y}, \mathbf{z}|\mathbf{x})} \right] + \log p_\theta(\mathbf{x}) \quad (4)$$

$$\begin{aligned} \mathcal{L}_{ELBO} &= \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})} [\log p_\theta(\mathbf{x}|\mathbf{z})] \\ &\quad - \mathbb{E}_{q_\phi(\mathbf{y}|\mathbf{x})} \mathcal{D} [q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y}) \parallel p_\theta(\mathbf{z}|\mathbf{y})] - \mathcal{D} [q_\phi(\mathbf{y}|\mathbf{x}) \parallel p(\mathbf{y})] \end{aligned} \quad (5)$$

4.1.4 ESTIMATION AND PROPAGATION

Sampling estimation for q_ϕ : TODO

Understanding the back-propagation for unsupervised case: $\mathbb{E}_{q_\phi(\mathbf{y}|\mathbf{x})} \mathcal{D} [q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y}) \parallel p_\theta(\mathbf{z}|\mathbf{y})]$

Lemma 1 *Latent prior model is a GMM with SGD update.*

Proof note: scaling gradient through propagation, $\sum_{\mathbf{y}}$, form of GMM model ■

Lemma 2 *GMVAE reduces to VaDE in simple case.*

Lemma 3 *SGD update is implicitly an MLE.*

4.2 Bayesian Repulsive Prior

find a nice drawing software for neural nets that allows depicting repulsion!

Lemma 4 *some lemma or theorem on covariance repulsion! Use HSIC*

4.3 Regularization

5. Experiments

5.1 Simulation

Simple MNIST is sufficient

5.2 Single Cell Clustering

Use The Human Cell Atlas data to show that 1. we can cluster them 2. with a simple linear transformation, those cells are mappable to their physical place! A cool picture ensues! [Weinstein et al. \(2013\)](#)

6. Conclusion

References

John N Weinstein, Eric A Collisson, Gordon B Mills, Kenna R Mills Shaw, Brad A Ozenberger, Kyle Ellrott, Ilya Shmulevich, Chris Sander, Joshua M Stuart, Cancer Genome Atlas Research Network, et al. The cancer genome atlas pan-cancer analysis project. *Nature genetics*, 45(10):1113, 2013.