Find your microenvironments faster with Neural Spatial LDA

Anonymous Author(s)
Affiliation
Address
email

Abstract

Spatial organization of different cell types in tissues have been shown to be important factors in many important biological processes such as aging, infection and cancer [3]. In particular, organization of the cells in a tumor microenvironment (TME) has been shown to play a crucial role in treatment response, disease pathology and patient outcome [14]. Spatial LDA [5] is a general purpose probabilistic model that has been used to discover novel microenvironments in settings such as Triple Negative Breast Cancer (TNBC) and Tuberculosis infections. However, the implementation of Spatial LDA proposed in [5] uses variational inference for learning model parameters and unfortunately does not scale well with dataset size and does not lend itself to speed-up via GPUs / TPUs. As researchers begin to collect larger in-situ multiplexed imaging datasets, there is a growing need for more scalable approaches for analysis of microenvironments. Here we propose a VAE-style network which we call *Neural Spatial LDA* extending the auto-encoding Variational Bayes formulation of classical LDA from [18]. We show Neural Spatial LDA achives significant speed-up over Spatial LDA while at the same time recovering similar topic distributions thus enabling its use in large data domains.

1 Introduction

The development of in-situ multiplexed profiling techniques, has enabled high-dimensional measurements of the abundance of protein & expression of genes while preserving spatial information [12, 4, 19, 17, 16, 11, 6]. For example, Keren *et al.* [8] characterized tumor-immune microenvironments in TNBC using Multiplixed Ion Bean Imaging by Time Of Flight (MIBI-ToF) [9] and categorized microenvironments into three types based on the spatial organization of tumor and immune cells and further demonstrated that these microenvironments stratified patient survival in their cohort. Similarly, Mccaffrey *et al.*[13] identified specific immuno-suppressive microenvironments in Tuberculosis granulomas which were associated with patient outcomes. Both examples demonstrate the potential for spatial analysis to yield prognostically relevant markers of disease.

While there exist various techniques to delineate cellular microenvironments [8, 15], in our work we focus on accelerating Spatial Latent Dirichlet Allocation (Spatial LDA) [5] which models cellular microenvironments in a probabilistic manner using topic models that incorporate a spatial prior - that neighboring microenvironments are more likely to have similar distribution of cell types. The current implementation of Spatial LDA uses mean field variational inference for determining the posterior distribution of the parameters of the generative model and does not leverage modern accelerators like GPUs or TPUs. As a result it takes hours to fit a model to even moderately sized datasets today.

In this work, we propose an encoder-decoder style network based on auto-encoding Variational Bayes LDA 18 with spatial regularization of discovered topics, which we call *Neural Spatial LDA*. Our implementation of Neural Spatial LDA runs on GPUs / TPUs and we show on several previously

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

published datasets that the microenvironments detected by Neural Spatial LDA are similar to Spatial LDA but with significantly faster training and inference times. Finally, reformulating Spatial LDA as Neural Spatial LDA also opens the possibility of further extensions to the Spatial LDA family of models, for example enabling supervised variants to specifically find microenvironments that are predictive of patient outcomes or regional annotations, extensions that we are currently exploring.

2 Neural Spatial LDA

As in Spatial LDA[5], we cast the identification of micro-environments from multiplexed images as a topic modeling problem, where a microenvironment is considered a *document*, unordered set of cells that belong to a microenvironment are analogous to *words in a document*. The task then is to identify both distribution cells / words per topic and distribution of *topics* for each microenvironment / document using cell counts. First, we briefly review LDA [2] and Spatial LDA, we then describe the auto-encoding Variational Bayes version of Spatial LDA. Spatial LDA assumes the data to have a generative process as follows,

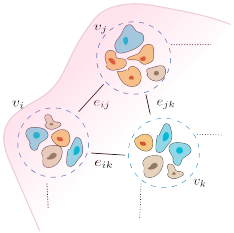

Figure 1: Illustration of microenvironments in a tissue as a graph.

$$
\begin{aligned}
p(\boldsymbol{\alpha}) &\propto \prod_{e_{ij} \in \mathcal{E}} Laplace(\alpha_i - \alpha_j, \sigma) \\
\boldsymbol{\theta}_k &\sim Dir(\alpha \mathbf{1}_K) \\
\boldsymbol{\beta}_v &\sim Dir(\eta \mathbf{1}_V) \\
\boldsymbol{z}_n &\sim Cat(\boldsymbol{\theta}_k) \\
\boldsymbol{w}_n &\sim Cat(\boldsymbol{\beta}_{\boldsymbol{z}_n})
\end{aligned}
$$

Let us consider each in-situ profiled image to be a graph $\mathcal{G}$ with vertices $\mathcal{V} = \{v_i, ...v_N\}$ representing microenvironments and edges $\mathcal{E}$ with each $e_{ij} \in \mathcal{E}$ denoting an edge from vertex v_i to v_j. Spatial LDA adds a Laplacian constraint over every pair of α_i & α_j of microenvironments v_i & v_j that are connected by an edge e_{ij} and introduces a hyperparameter σ controling the *strength* of this prior over topic distributions of microenvironments. Intuitively, the graph $\mathcal{G}$ defines neighborhoods that should have similar topic distributions and σ controls how strongly that inductive bias is enforced. The rest of the generative process is a standard LDA model with each cell $\boldsymbol{w}_n$ being sampled by first choosing a topic distribution $\boldsymbol{\theta}_k$ from a Dirichlet prior parameterized by α, then choosing a topic $\boldsymbol{z}_n$ and then further choosing a cell from the categorical distribution of cells parameterized by $\boldsymbol{\beta}_{\boldsymbol{z}_n}$. Here, K represents the number of topics and V represents the size of the vocabulary.

Note that, in the limit where the graph $\mathcal{G}$ is a fully connected graph (i.e.) $\mathcal{E} = \{e_{ij} \mid \forall (i,j) \in \{1, ..., N\}, i \neq j\}$ and $\sigma \to 0$, all α's are constrained to be the same and Spatial LDA reduces to classical LDA. In that scenario, all microenvironments are assumed to have the same cell distributions and spatial information is lost. In order to capture spatial structure in multiplexed images Spatial LDA defines a graph that captures spatial relationships between local neighborhood of cells. In the original paper, the authors suggest defining $\mathcal{E}$ by taking the Voronoi partitioning of neighborhood positions and connecting neighborhoods that share a facet in Voronoi partitions.

Spatial LDA then uses mean field variational inference for learning the parameters of the model by optimizing the Evidence Lower BOund (ELBO) shown in (1) where γ, ϕ, ξ are variational parameters for approximating distributions parametrized by $\boldsymbol{\theta}, \boldsymbol{z}, \boldsymbol{\alpha}$ respectively.

$$
\mathcal{L}(\gamma, \phi, \xi | \sigma, \beta) = -D_{KL}[q(\boldsymbol{\theta}, \boldsymbol{z}, \boldsymbol{\alpha} | \gamma, \phi, \xi) || p(\boldsymbol{\theta}, \boldsymbol{z}, \boldsymbol{\alpha} | \sigma)] + \mathbb{E}_{q(\boldsymbol{\theta}, \boldsymbol{z}, \boldsymbol{\alpha} | \gamma, \phi, \xi)}[\log p(\boldsymbol{w} | \boldsymbol{\theta}, \boldsymbol{z}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma)] \quad (1)
$$

Owing to the conjugacy of categorical and Dirichlet distributions, the coordinate update steps have a closed form in classical LDA, while Spatial LDA involves optimizing a non-smooth function for the variational parameter ξ, due to the presence of the Laplacian prior. This was overcome by

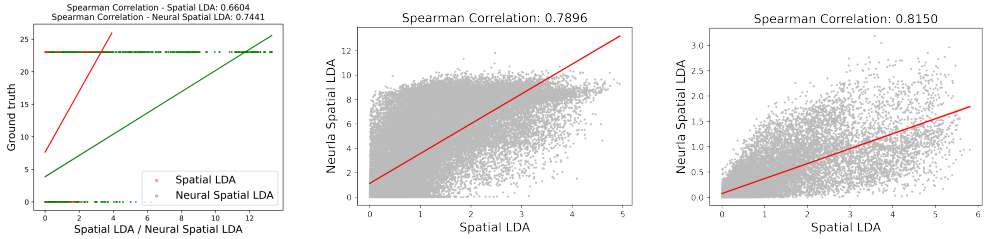

Figure 2: Comparison of KL-divergence of topic distributions between pairs of cells inferred using Spatial LDA and Neural Spatial LDA for Simulated (*left*) using 50k random pairs, Mouse spleen (*center*) using 50k random pairs & TB (*right*) using 10k random pairs.

using alternating direction method of multipliers (ADMM) + primal-dual interior point optimization approaches (we refer the reader to Spatial LDA[5] for the exact update steps). We observe that these optimization steps do not scale well with the number of data points compared to gradient based methods for training modern deep neural networks. In addition, parameter updates of Spatial LDA are not optimized to run on GPUs / TPUs. Therefore, to enable faster training and inference, we propose to use auto-encoding Variational Bayes (AEVB)[10] for Spatial LDA.

Srivastava *et al.*[18] proposed two key ingredients that made it possible to apply the re-parametrization trick to classical LDA. Specifically, 1) Marginalize out the discrete variables $\boldsymbol{z}$'s 2) Replace Dirichlet prior with Logistic Normal with diagonal covariance in softmax basis. Logistic normal could be easily re-parametrized by sampling from the standard normal as, $\mathbb{E}_{\boldsymbol{z}\sim\mathcal{LN}(\boldsymbol{\mu},\boldsymbol{\Sigma})}[f(\boldsymbol{z})] = \mathbb{E}_{\boldsymbol{\epsilon}\sim\mathcal{N}(0,I)}[f(\rho(\boldsymbol{\mu} + \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\epsilon}))]$ where ρ represents the softmax function. Neural Spatial LDA uses these two ideas along with spatial regularization as given below,

2.1 Spatial Regularization

The final piece in the puzzle and our main contribution is to enforce spatial regularization on topic distribution of microenvironments. This could be done through two different ways, 1) sample α's such that they obey $p(\boldsymbol{\alpha}) \propto \prod_{e_{ij}\in\mathcal{E}} Laplace(\alpha_i - \alpha_j, \sigma)$ and compute prior $p(\boldsymbol{\theta}|\boldsymbol{\alpha}, \sigma) = \mathcal{LN}(\boldsymbol{\theta}|\boldsymbol{\alpha}, \sigma; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ that will impose the constraint through the KL-term of the ELBO or 2) apply the Laplacian constraint directly on $\boldsymbol{\mu}$'s. Since VAEs learn variational posteriors are far away from the original priors they start with after training, we opted to apply L1 constraint on $\boldsymbol{\mu}$'s of microenvironments that are connected via an edge in $\mathcal{G}$. The final objective function for our proposed Neural Spatial LDA is the following,

$$\mathcal{L}(\gamma, \phi|\sigma, \beta) = -D_{KL}[q_\gamma(\boldsymbol{\mu}, \boldsymbol{\Sigma}|\boldsymbol{\alpha})||p(\boldsymbol{\mu}, \boldsymbol{\Sigma}|\boldsymbol{\alpha})] + \mathbb{E}_{q_\gamma(\boldsymbol{\mu},\boldsymbol{\Sigma}|\boldsymbol{\alpha})}[\log p(\boldsymbol{w}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma)] + \frac{1}{|\mathcal{E}|}\sum_{e_{ij}\in\mathcal{E}} \frac{1}{\sigma}|\boldsymbol{\mu_i} - \boldsymbol{\mu_j}|$$

3 Experiments

We applied Neural Spatial LDA to three different datasets - 1) *Simulated:* A simulated dataset with known ground truth topic distribution and "cell" distribution under each topic (details in Appendix). 2) *Mouse spleen:* A dataset of CODEX images of mouse spleen with a panel of 30 different antibodies from [7]. 3) *TB:* A dataset of MIBI images of granulomas in Tuberculosis patients from [13].

As in the original Spatial LDA paper, we preprocess CODEX and MIBI-ToF multichannel images to segment cells using DeepCell [1]. Then we get mean expression profiles of each cell and then cluster these profiles using FlowSOM [20]. We then associate cell index to each cluster and then get the count of different types of cells in a spatial neighborhood of each cell. The cell counts of different cell types around a spatial neighborhood constitute a microenvironment and we use them as input features.

To qualitatively evaluate whether Spatial LDA and Neural Spatial LDA infer similar topic distributions to various microenvironments, we run Spatial LDA, & Neural Spatial LDA on all three datasets and

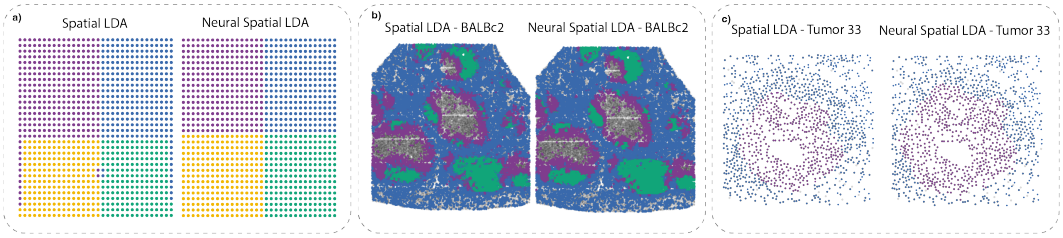

Figure 3: Qualitative comparison of the most probable topic for each cell between Spatial LDA and Neural Spatial LDA on (*a*) Simulated data (*b*) Mouse spleen (*c*) Tuberculosis data with number of topics set to three.

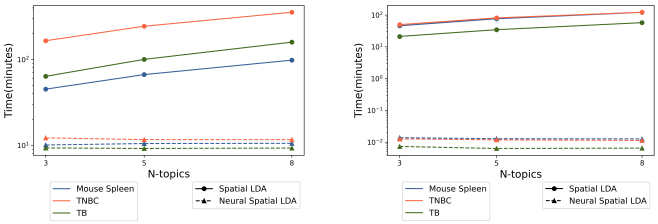

Figure 4: Training (*Left*) & Inference (*Right*) time difference between Spatial LDA and Neural Spatial LDA on all three datasets for different number of topics.

color each cell by its most probable topic. From Fig.[3], we observe that Spatial LDA and Neural Spatial LDA groups similar regions of cells into similar topics across all three datasets. (note: the topics are matched manually after training since the order of the topics is not guaranteed to match)

To quantitatively measure the similarity between topic *distributions* learnt by both methods, we randomly sample pairs of cells from each image without replacement, compute the KL-divergence between the topic distributions of pairs of cells and report the Spearman correlation between Spatial LDA and Neural Spatial LDA. Our intuition is that this KL divergence should capture the similarity between 2 neighborhoods under each model and a high correlation indicates that Spatial and Neural Spatial LDA both capture similar notions of similarity / differences in neighborhoods. This also allows us to evaluate the agreement between Spatial and Neural Spatial LDA in a way that is agnostic to reordering of topics. We see from Fig.[2], that Spatial LDA and Neural Spatial LDA yields highly correlated KL-divergences between topic distributions for the same pairs of cells indicating good agreement between the two methods. We note that we expect some differences in KL-divergence due to the approximations described above - marginalizing $z's$ & replacing Dirichlet with Logistic Normal, as well as differences - in optimization co-ordinate ascent vs gradient descent - used for inference.

Next, we measure the time it takes for Spatial LDA and Neural Spatial LDA to train using the same machine, with Spatial LDA was run on the CPU as it was not optimized to run on GPUs while Neural Spatial LDA was run a single Titan RTX. The stopping criteria for Spatial LDA is set to be the same number of iterations used in the original work while Neural Spatial LDA was trained til convergence of the ELBO. From Fig.[4], we could see that AEVB version of Spatial LDA achieves $\sim$84%, $\sim$95% & $\sim$90% reduction in training times on Mouse spleen, TNBC & Tuberculosis datasets respectively.

4 Conclusion

Spatial LDA is a probabilistic model used to identify TMEs in an unsupervised way but its inability to utilize GPUs / TPUs and slow training and inference limits its use in large data domain applications. We propose Neural Spatial LDA and show on previously published datasets that we recover similar microenvironments while acclerating training and inference significantly.

References

Dylan Bannon, Erick Moen, Enrico Borba, Andrew Ho, Isabella Camplisson, Brian Chang, Eric Osterman, William Graf, and David Van Valen. Deepcell 2.0: Automated cloud deployment of deep learning models for large-scale cellular image analysis. 2018.

David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

Katie E Blise, Shamilene Sivagnanam, Grace L Banik, Lisa M Coussens, and Jeremy Goecks. Single-cell spatial architectures associated with clinical outcome in head and neck squamous cell carcinoma. *NPJ precision oncology*, 6(1):1–14, 2022.

Kok Hao Chen, Alistair N Boettiger, Jeffrey R Moffitt, Siyuan Wang, and Xiaowei Zhuang. Spatially resolved, highly multiplexed rna profiling in single cells. *Science*, 348(6233):aaa6090, 2015.

Zhenghao Chen, Ilya Soifer, Hugo Hilton, Leeat Keren, and Vladimir Jojic. Modeling multiplexed images with spatial-lda reveals novel tissue microenvironments. *Journal of Computational Biology*, 27(8):1204–1218, 2020.

Simone Codeluppi, Lars E Borm, Amit Zeisel, Gioele La Manno, Josina A van Lunteren, Camilla I Svensson, and Sten Linnarsson. Spatial organization of the somatosensory cortex revealed by osmfish. *Nature methods*, 15(11):932–935, 2018.

Yury Goltsev, Nikolay Samusik, Julia Kennedy-Darling, Salil Bhate, Matthew Hale, Gustavo Vazquez, Sarah Black, and Garry P Nolan. Deep profiling of mouse splenic architecture with codex multiplexed imaging. *Cell*, 174(4):968–981, 2018.

Leeat Keren, Marc Bosse, Diana Marquez, Roshan Angoshtari, Samir Jain, Sushama Varma, Soo-Ryum Yang, Allison Kurian, David Van Valen, Robert West, et al. A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell*, 174(6):1373–1387, 2018.

Leeat Keren, Marc Bosse, Steve Thompson, Tyler Risom, Kausalia Vijayaragavan, Erin McCaffrey, Diana Marquez, Roshan Angoshtari, Noah F Greenwald, Harris Fienberg, et al. Mibi-tof: A multiplexed imaging platform relates cellular phenotypes and tissue structure. *Science advances*, 5 (10):eaax5851, 2019.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Jia-Ren Lin, Benjamin Izar, Shu Wang, Clarence Yapp, Shaolin Mei, Parin M Shah, Sandro Santagata, and Peter K Sorger. Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-cycif and conventional optical microscopes. *Elife*, 7, 2018.

Eric Lubeck, Ahmet F Coskun, Timur Zhiyentayev, Mubhij Ahmad, and Long Cai. Single-cell in situ rna profiling by sequential hybridization. *Nature methods*, 11(4):360–361, 2014.

Erin F McCaffrey, Michele Donato, Leeat Keren, Zhenghao Chen, Alea Delmastro, Megan B Fitzpatrick, Sanjana Gupta, Noah F Greenwald, Alex Baranski, William Graf, et al. The immunoregulatory landscape of human tuberculosis granulomas. *Nature immunology*, 23(2):318–329, 2022.

Jeffrey R Moffitt, Emma Lundberg, and Holger Heyn. The emerging landscape of spatial profiling technologies. *Nature Reviews Genetics*, pages 1–19, 2022.

Tyler Risom, David R Glass, Inna Averbukh, Candace C Liu, Alex Baranski, Adam Kagel, Erin F McCaffrey, Noah F Greenwald, Belén Rivero-Gutiérrez, Siri H Strand, et al. Transition to invasive breast cancer is associated with progressive changes in the structure and composition of tumor stroma. *Cell*, 185(2):299–310, 2022.

Samuel G Rodriques, Robert R Stickels, Aleksandrina Goeva, Carly A Martin, Evan Murray, Charles R Vanderburg, Joshua Welch, Linlin M Chen, Fei Chen, and Evan Z Macosko. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science*, 363(6434):1463–1467, 2019.

Sheel Shah, Eric Lubeck, Wen Zhou, and Long Cai. In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron*, 92(2):342–357, 2016.

Akash Srivastava and Charles Sutton. Autoencoding variational inference for topic models. *arXiv preprint arXiv:1703.01488*, 2017.

Patrik L Ståhl, Fredrik Salmén, Sanja Vickovic, Anna Lundmark, José Fernández Navarro, Jens Magnusson, Stefania Giacomello, Michaela Asp, Jakub O Westholm, Mikael Huss, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353(6294): 78–82, 2016.

Sofie Van Gassen, Britt Callebaut, Mary J Van Helden, Bart N Lambrecht, Piet Demeester, Tom Dhaene, and Yvan Saeys. Flowsom: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry Part A*, 87(7):636–645, 2015.

A Simulated Dataset

We create a simulated dataset according to the $p(\boldsymbol{\theta})$ and $p(\boldsymbol{\beta})$ shown in (2) & (3), given (x, y) co-ordinate of a "cell".

$$p(\boldsymbol{\theta}|x,y) = \begin{cases} Cat & [1., 0., 0., 0.] \quad \text{if } x \leq H//2, y \leq W//2 \\ Cat & [0., 1., 0., 0.] \quad \text{if } x > H//2, y \leq W//2 \\ Cat & [0., 0., 1., 0.] \quad \text{if } x \leq H//2, y > W//2 \\ Cat & [0., 0., 0., 1.] \quad \text{if } x > H//2, y > W//2 \end{cases} \tag{2}$$

$$p(\boldsymbol{\beta}|\boldsymbol{z}) = \begin{bmatrix} 1. & \dots & 0. \\ & \ddots & \\ 0. & \dots & 1. \end{bmatrix} \in \mathcal{R}^{4\times|V|} \tag{3}$$

We set $H = 32$, $W = 32$ & $V = 4$ for our experiments on this simulated dataset. Fig.[5] shows the topic distribution for a 32×32 grid of cells.

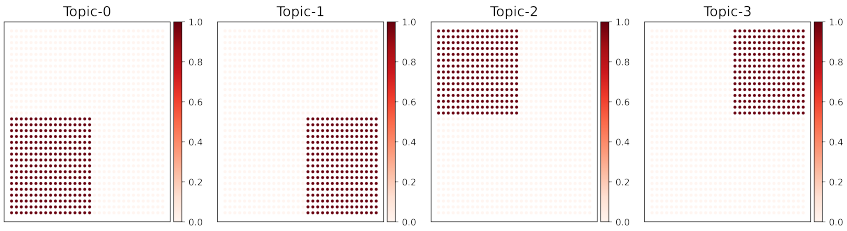

Figure 5: Topic distribution for the simulated dataset