

# Generative Models for SpatioTemporal Topic Modeling

Newt Gingrich

October 11, 2018

## 1 AR Process on log Dirichlet Concentration

### Notation:

Indices:

- $K$  number of topics
- $V$  size of vocabulary
- $L$  number of locations
- $Y$  number of time points
- $P$  number of spatial dimensions

Hyperparameters:

- $\eta$  dirichlet hyperparameter for word-topic concentrations

Parameters:

- $\Omega \in \mathbb{R}^{Y \times L \times K}$  log concentration parameters for document-topic prevalence.
- $\sigma_\omega$  isotropic error for matrix AR process on  $\Omega$ .
- $\phi \in \mathbb{R}^{K \times V}$  topic - word probability matrix.
- $w_{y,l,m,n}$  the word in position  $n$  of the  $m$ 'th document at location  $l$  at time  $y$ .
- $z_{y,l,m,n}$  the topic associated with  $w_{y,l,m,n}$ .
- $\mathbf{B}^{L \times L}$  matrix based on distances between the spatial locations divided by a scalar length scale
- $\mathbf{A}^{Y \times L \times K}$  matrix based on an AR(1) process

---

**Algorithm 1** StLda1

---

```
1:                                     ▷ Generate Spatiotemporal priors
2:  $\Omega_1 \sim N(\mathbf{0}, \sigma_\omega \mathbf{I})$ 
3: for  $y \in \{2, \dots, Y\}$  do
4:    $\epsilon_t \sim \text{iid} N(0, \sigma_\omega)$ 
5:    $\Omega_y \leftarrow \mathbf{B}\Omega_{y-1} + \epsilon_t$ 
6:                                     ▷ Generate Topics (standard LDA)
7: for  $k \in \{1, \dots, K\}$  do
8:    $\phi_k \sim \text{Dir}(\eta)$ 
9:                                     ▷ Generate Document-Topic Prevalences
10: for  $y \in \{1, \dots, Y\}$  do
11:   for  $l \in \{1, \dots, L\}$  do
12:     for  $m \in \{1, \dots, M\}$  do
13:        $\theta_{y,l,m} \sim \text{softmax}[\mathbf{N}(\Omega_{y,l}, \sigma_\theta)]$ 
14:                                     ▷ Generate Document-Topic Prevalences
15: for  $y \in \{1, \dots, Y\}$  do
16:   for  $l \in \{1, \dots, L\}$  do
17:     for  $m \in \{1, \dots, M\}$  do
18:       for  $n \in \{1, \dots, N_m\}$  do
19:          $z_{y,l,m,n} \sim \text{Mult}(\theta_{y,l,m})$ 
20:          $w_{y,l,m,n} \sim \text{Mult}(\phi_{z_{y,l,m,n}})$ 
```

---