

Scribe Notes: Bayesian Factorization of Big Sparse Tensors

Maja R Rudolph

March 13, 2016

Last weeks:

Word embeddings from a probabilistic modelling perspective

Now:

Statistics literature (no graphical models). We will draw the graphical models.

Mixture Model:

$$y_i = \{y_{i1}, \dots, y_{ip}\}$$

example applications:

- **genetics:** allele of person i at location j
- **survey:** response of person i on question j

Dunson and Xing 2009 [2]:

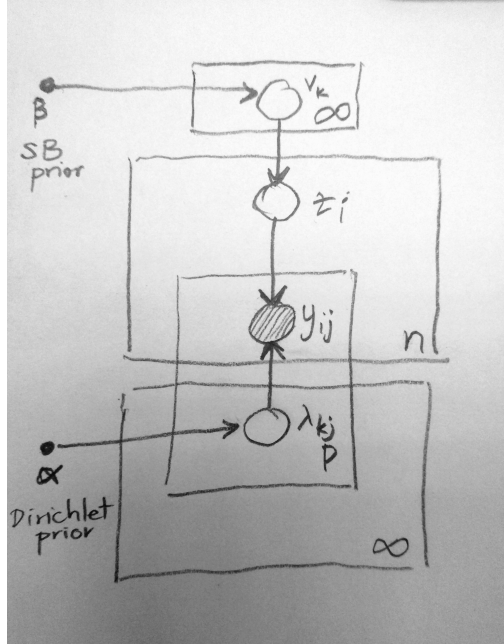
"Product of Multinomial Model"

$$\pi_{c_1 \dots c_p} = p(y_i | z_i) = \prod_{j=1}^p \lambda_{z_i, j}^{y_{ij}}$$

where

$$\lambda_{kj}^c := p(y_i = c | z_j = k)$$

The model captures trade-off between latent class vs. background distribution, which doesn't differ person by person but question by question.



Stickbreaking Prior (SB):

$$\pi \sim \text{SB}(\beta)$$

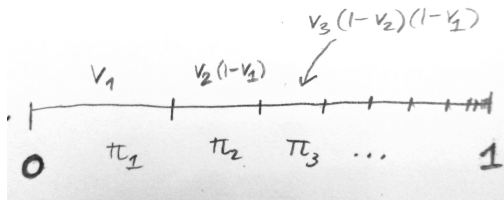
$$z_i \sim \pi$$

$$y_i \sim p(y_i | z_i)$$

where for scalar β

$$v_k \sim \text{Beta}(1, \beta)$$

$$\pi_k = v_k \prod_{i=1}^{k-1} (1 - v_i)$$



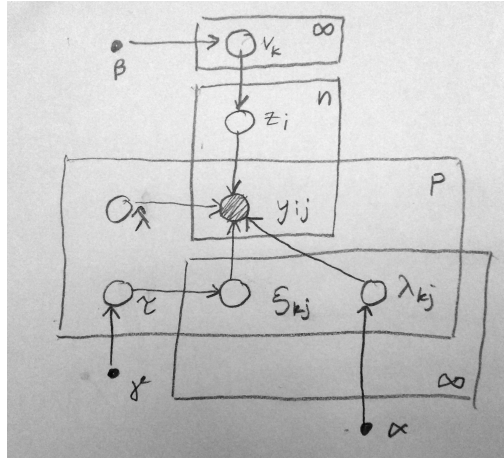
This prior let's us define an infinite mixture model.

Connection to dirichlet process $DP(\alpha G_0)$ as discrete random measure. (Any permutation of the space).

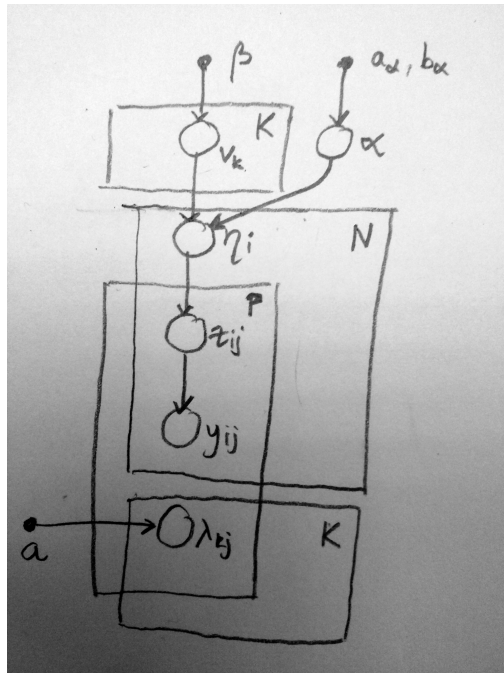
$\sum_{i=1}^{\infty} \pi_i \delta_{\theta_i^*}$ relates to $\pi \sim \text{SB}(\beta)$. For more details see Sethuraman 1994 [3].

The size-biasedness of π might help resolve identifiability issues.

Zhou, Bhattacharya, Herring and Dunson 2015 [4]:



Bhattacharya and Dunson 2012 [1]:



"Mixed Membership Model"

Posterior Computation: Prior is nonconjugate. Need to do Metropolis Hastings for α

References

- [1] Bhattacharya, A. and Dunson, D. B. (2012). Simplex factor models for multivariate unordered categorical data. *Journal of the American Statistical Association*, 107(497):362–377.
- [2] Dunson, D. B. and Xing, C. (2009). Nonparametric bayes modeling of multivariate categorical data. *Journal of the American Statistical Association*, 104(487):1042–1051.
- [3] Sethuraman, J. (1994). A constructive definition of dirichlet priors. *Statistica sinica*, pages 639–650.
- [4] Zhou, J., Bhattacharya, A., Herring, A. H., and Dunson, D. B. (2015). Bayesian factorizations of big sparse tensors. *Journal of the American Statistical Association*, 110(512):1562–1576.