Supplement to Joint Modeling Matrix and Texts with Latent Binary Features

XianXing Zhang
Duke University
xianxing.zhang@duke.edu

Lawrence Carin
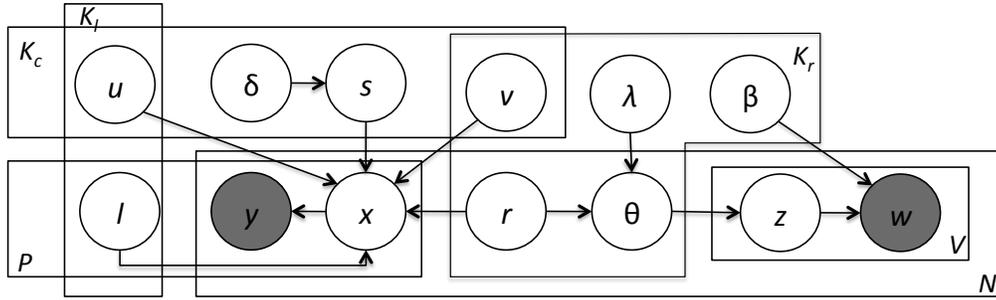Duke University
lcarin@ee.duke.edu

1 Graphical Model

Figure 1: *Graphical representation of the proposed model (with the hyperparameters omitted).*

2 Notation

In the following we denote $l_{i:}u_{:k} = a_{ik}$, $r_{j:}v_{:k} = b_{jk}$. By noting that binary vectors $l_{i:}$ and $r_{j:}$ have only finite of "1"s (as a consequence of the property of IBP), and both $l_{i:}$ and $r_{j:}$ have bounded variance, it's straightforward to show that $\mathbb{E}(a_{ik})^2$ and $\mathbb{E}(b_{jk})^2$ are bounded by constant $a < \infty$ and $b < \infty$ respectively.

3 Proof of Theorem 1

To prove the sequence $\{\sum_{k=1}^{K} s_k a_{ik} b_{jk}\}_{K>0}$ converges in ℓ_2 we make use of the following lemma:

Lemma 1. *(Cauchy criterion) A sequence of random variables $\{X_n\}_{n \geq 1}(\mathbb{E}X_n^2 < \infty)$ converges to a random variable X $(\mathbb{E}X^2 < \infty)$ in ℓ_2 iff $\mathbb{E}(X_{n+k} - X_k)^2 \to 0$ as $n, k \to \infty$.*

Based on Lemma 1, to prove Theorem 1 it's sufficient to show that $\lim_{k \to \infty} \mathbb{E}(\sum_{k=h}^{\infty} s_k a_{ik} b_{jk})^2 = 0$.

Based on the way we constructed the model, s_k and a_{ik}, b_{jk} are independent, and $\mathbb{E}(s_k) = 0$, thus we have

$$\mathbb{E}\left(\sum_{k=h}^{\infty} s_h a_{ih} b_{jh}\right)^2 = \sum_{k=h}^{\infty} \mathbb{E}(s_h)^2 \mathbb{E}(a_{ih})^2 \mathbb{E}(b_{jh})^2$$

where $\mathbb{E}(s_h)^2 = \mathbb{E}(\mathbb{E}((s_h)^2|\tau_h)) = \mathbb{E}\left(\frac{1}{\prod_{l=1}^{h} \delta_l}\right) = \frac{1}{\alpha_c^h}$, where $\mathbb{E}(\delta_l) = \alpha_c > 1$. Substitute this into the above equation we obtain

$$\sum_{k=h}^{\infty} \mathbb{E}(s_h)^2 \mathbb{E}(a_{ih})^2 \mathbb{E}(b_{jh})^2 \leq \sum_{k=h}^{\infty} \frac{ab}{\alpha_c^h} = \frac{ab(1 - 1/\alpha_c)}{\alpha_c^{k-1}}$$

Thus by letting $k \to \infty$ we arrive at the desired result

$$\lim_{k\to\infty} \mathbb{E}\left(\sum_{k=h}^{\infty} s_h a_{ih} b_{jh}\right)^2 = \lim_{k\to\infty} \frac{ab(1 - 1/\alpha_c)}{\alpha_c^{k-1}} = 0$$

Thus sequence $\{\sum_{k=1}^{K} s_k a_{ik} b_{jk}\}_{K>0}$ converges in ℓ_2 when $\alpha_c > 1$.

4 Proof of Theorem 2

The proof of Theorem 2 is similar to its counterpart in [1]. Denote $N_{ij}^K = |M_{ij}^K - M_{ij}^\infty|^2 = \left(\sum_{k=K+1}^{\infty} s_k a_{ik} b_{jk}\right)^2$ we have

$$p\{N_{ij}^K \leq \epsilon\} = E\{p(N_{ij}^K \leq \epsilon|\boldsymbol{\tau})\} = 1 - E\{p(N_{ij}^K > \epsilon|\boldsymbol{\tau})\} > 1 - \frac{E\left(E(N_{ij}^K|\boldsymbol{\tau})\right)}{\epsilon}$$

where $\frac{E\left(E(N_{ij}^K|\boldsymbol{\tau})\right)}{\epsilon} = \frac{\mathbb{E}(a_{iK+1})^2 \mathbb{E}(b_{jK+1})^2(1-1/\alpha_c)}{\alpha_c^K} \leq \frac{ab(1-1/\alpha_c)}{\alpha_c^K}$ as we showed above. Thus we have

$$p\{|M_{ij}^K - M_{ij}^\infty|^2 > \epsilon\} < \frac{ab(1 - 1/\alpha_c)}{\epsilon \alpha_c^K}$$

References

[1] A. Bhattacharya and D. B. Dunson. Sparse Bayesian infinite factor models. *Biometrika*, 2011.