Bayesian Boolean Matrix Factorisation – Supplementary Information

Tammo Rukat [1] **Chris C. Holmes** [1 2] **Michalis K. Titsias** [3] **Christopher Yau** [4 5]

A. Derivation of the Conditionals

Here we derive the full conditional for a latent variable z_{nl} as stated in eq. (4). The expression for u_{ld} is analogous. We start from the full joint:

$$p(\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{Z}, \lambda) = p(\boldsymbol{X}|\boldsymbol{U}, \boldsymbol{Z}, \lambda)p(\boldsymbol{Z})p(\boldsymbol{U})p(\lambda) \ .$$

We drop terms that do not depend on z_{nl}, and plug in the likelihood from eq. (3):

$$p(z_{nl}|\cdot) \propto \prod_d \sigma \left[\lambda \tilde{x}_{nd}(1{-}2 \prod_{l'} (1{-}z_{nl'}u_{l'd})) \right] p(z_{nl}) \ .$$

Next, we normalise this expression for $z_{nl} \in [0, 1]$ and restrict to $z_{nl} = 1$ for simplicity:

$$p(z_{nl} = 1|\cdot) = \sigma \left[\text{logit}(p(z_{nl}) + \sum_d \log \frac{1{+}\exp[-\lambda \tilde{x}_{nd} \, (1{-}2 \prod_{l'} (1{-}z_{nl'}u_{l'd}))]_{z_{nl}=0}}{1{+}\exp[-\lambda \tilde{x}_{nd} \, (1{-}2 \prod_{l'} (1{-}z_{nl'}u_{l'd}))]_{z_{nl}=1}} \right] . \tag{9}$$

We can distinctly simplify the second line of equation eq. (9) by distinguishing the two possible contributions to the sum. Changing z_{nl} from 0 to 1 in the fraction inside the sum can have the following two consequences:

1. Numerator and denominator remain equal, then the contribution to the sum is zero.
2. The numerator's exponent evaluates to $\lambda \tilde{x}_{nd}$ and the denominator's exponent to $-\lambda \tilde{x}_{nd}$, then the contribution to the sum is $\lambda \tilde{x}_{nd}$, as we can see by using the identity $\log(1 + e^x) - \log(1 + e^{-x}) = x$.

For scenario 2 to take place, we can see from eq. (9) that two conditions need to be met:

1. $u_{ld} = 1$. Otherwise, the value z_{nl} does not effect the likelihood. Viewed es directed graphical model, there would be no link between z_{nl} and x_{nd}.
2. $z_{n'l}u_{l'd} = 0 \ \forall \ l' \neq l$. Otherwise, another parent already explains x_{nd}. Viewed as directed graphical model, x_{nd} would be *explained away*.

Thus we find the equality for the fraction in eq. (9):

$$\sum_d \log \frac{1{+}\exp[-\lambda \tilde{x}_{nd} \, (1{-}2 \prod_{l'} (1{-}z_{nl'}u_{l'd}))]_{z_{nl}=0}}{1{+}\exp[-\lambda \tilde{x}_{nd} \, (1{-}2 \prod_{l'} (1{-}z_{nl'}u_{l'd}))]_{z_{nl}=1}} = \lambda \sum_d \tilde{x}_{nd} u_{ld} \prod_{l' \neq l} (1{-}z_{nl'}u_{l'd}) \ .$$

This leads to the full conditional as given in eq. (4):

$$p(z_{nl}|\cdot) = \sigma \left[\text{logit}(p(z_{nl})){+}\lambda \tilde{z}_{nl} \sum_d \tilde{x}_{nd} u_{ld} \prod_{l' \neq l} (1{-}z_{nl'}u_{l'd}) \right] \ .$$

[1]Department of Statistics, University of Oxford, UK [2]Nuffield Department of Medicine, University of Oxford, UK [3]Department of Informatics, Athens University of Economics and Business, Greece [4]Centre for Computational Biology, Institute of Cancer and Genomic Sciences, University of Birmingham, UK [5]Wellcome Trust Centre for Human Genetics, University of Oxford, UK. Correspondence to: Tammo Rukat <tammo.rukat@stats.ox.ac.uk>.

Proceedings of the 34^{th} International Conference on Machine Learning, Sydney, Australia, PMLR 70, 2017. Copyright 2017 by the author(s).