SKETCH OF THE "ALTERNATING SQP" METHOD FOR FITTING POISSON TOPIC MODELS

PETER CARBONETTO*

1. Some derivations. Given an $n \times p$ matrix of counts X with individual entries $x_{ij} \geq 0$, our aim is to fit a Poisson model of the counts,

$$\begin{aligned} p(x) &= \prod_{i=1}^n \prod_{j=1}^p p(x_{ij}) \\ &= \prod_{i=1}^n \prod_{j=1}^p \mathrm{Poisson}(x_{ij}; \lambda_{ij}), \end{aligned} \tag{1.1}$$

in which the Poisson rates are given by the mixture $\lambda_{ij} = \sum_{k=1}^K l_{ik} f_{jk}$. Therefore, the Poisson model is specified by a $p \times K$ matrix F with entries $f_{ik} \geq 0$ (the "factors"), and an $n \times K$ matrix L with entries $l_{ik} \geq 0$ (the "loadings"). Fitting F and L is equivalent to non-negative matrix factorization with the "beta-divergence" cost function [2]. It can also be used to recover a maximum-likelihood estimate for the latent Dirichlet allocation (LDA) model [1]. So fitting this model is useful for a wide range of applications.

The log-likelihood for the Poisson model is

$$\log p(x \mid F, L) \propto \sum_{i=1}^n \sum_{j=1}^p x_{ij} \log(\textstyle\sum_{k=1}^K l_{ik} f_{jk}) - \displaystyle\sum_{i=1}^n \sum_{j=1}^p \sum_{k=1}^K l_{ik} f_{jk}, \tag{1.2}$$

where the constant of proportionality is obtained from factorial terms in the Poisson densities. Our specific aim is to find a F and L that maximizes the log-likelihood. In the remainder, we derive an efficient approach to doing this.

REFERENCES

[1] D. M. Blei, A. Y. Ng, and M. I. Jordan, *Latent Dirichlet allocation*, Journal of Machine Learning Research, 3 (2003), pp. 993–1022.

[2] D. D. Lee and H. S. Seung, *Algorithms for non-negative matrix factorization*, in Advances in Neural Information Processing Systems 13, 2001, pp. 556–562.

*Dept. of Human Genetics and the Research Computing Center, University of Chicago, Chicago, IL