Poisson Factor Analysis

Huiqiang

November 11, 2019

Abstract
This code is the implementation of basic PFA model by Gibbs Sampling.

1 Train and Test data

1. The data is a $N * 3$ matrix.
2. The first column is the document index.
3. The second column is the word index.
4. The third column is the number of word in document.

2 Beta-Gamma-Gamma-Poisson Model

$$x_{pi} = \sum_{k=1}^{K} x_{pik}, x_{pik} \sim Pois(\phi_{pk}\theta_{ki}) \tag{1}$$
$$\phi_k \sim Dir(\alpha, \dots \alpha) \tag{2}$$
$$\theta_{ki} \sim Gamma(r_k, \frac{p_k}{1-p_k}) \tag{3}$$
$$r_k \sim Gamma(c_0 * r_0, 1/c_0) \tag{4}$$
$$p_k \sim Beta(c\epsilon, c(1-\epsilon)) \tag{5}$$

- x_{pi} is the count of term p in document i.
- ϕ_k is topic-work matrix.

3 MCMC Inference

1. Initialization of Hyperparameters.
2. Sample x_{pik} [1]

$$x_{pik} \sim Mult(x_{pi}; \zeta_{pik}) \tag{6}$$
$$\zeta_{pik} = \frac{\phi_{pk}\theta_{ki}}{\sum_{k=1} K\phi_{pk}\theta_{ki}} \tag{7}$$

3. Sampling ϕ_k.

$$x_{.ik} = \sum_{p=1}^{P} x_{pik} \tag{8}$$
$$x_{.ik} = Pois(\sum_{p=0}^{P} \phi_{pk}\theta_{ki}) \tag{9}$$
$$\sum_{p=0}^{P} \phi_{pk} = 1 \tag{10}$$
$$p(x_{1ik}, \dots x_{pik}) \sim Mult(x_{.ik}; \phi_k) \tag{11}$$
$$p(\phi_k|-) \sim Dir(\alpha + x_{p.k})\ Given\ Equation(2) \tag{12}$$

4. Sampling p_k
   Beta Distribution is the conjugate prior of Negative Binomial Distribution.Marginalizng ϕ_k *and* θ_{ki} *out*.

$$
\begin{aligned}
x_{.ik} &\sim NB(r_k, p_k) && (13)\\
p_k &\sim Beta(c\epsilon\, c(1-\epsilon)) && (14)\\
f(p_k|x_{..k}) &\propto f(x_{..k}|p_k) f(p_k) && (15)\\
f(x_{..k}|p_k) &= \prod_{i=1}^{doc} f(x_{.ik}|p_k) && (16)\\
f(p_k)|-) &\sim Beta(c(1-\epsilon) + x_{..k}, c(1-\epsilon) + N * r_k) && (17)
\end{aligned}
$$

5. Sampling r_k[2]

$$
\begin{aligned}
p(r_k|-) &\propto Gamma(r_k; c_0 r_0, 1/c_0) \prod_{i=1}^{doc} NB(x_{.ik}; r_k, p_k) && (18)\\
x_{.ik} &\sim \sum_{t=1}^{l_k} \log(p_k), l_{ik} \sim Pois(-r_k \log(1-p_k)) && (19)\\
l_{ik} &\sim CRT(x_{.ik}, p_k), x_{.ik} \sim NB(r_k, p_k) && (20)
\end{aligned}
$$

   Then we can get l_{ik}.By equation (19) and l_{ik}.

$$
p(r_k|-) \sim Gamma(c_0 r_0 + \sum_{i=1}^{doc} l_{ik}, \frac{1}{c_0 - N\log(1-p_k)}) \qquad (21)
$$

6. Sampling θ_{ki}.

$$
\begin{aligned}
x.ik &= \sum_{p=1}^{voc} x_{pik} \sim Pois(\sum_{p=1}^{voc} \phi_{pk}\theta_{ki}) = Pois(\theta_{ki}) && (22)\\
\theta_{ki} &\sim Gamma(r_k, \frac{p_k}{1-p_k}) && (23)\\
f(_{ki}|-) &\sim Gamma(r_k + x_{.ik}, p_k) && (24)
\end{aligned}
$$

7. Compute Perplexity

$$
\begin{aligned}
\lambda_{pi} &= \frac{\sum_k \phi_{pk}\theta_{ki}}{\sum_k \sum_p \phi_{pk}\theta_{ki}} && (25)\\
Perplexity &= \exp(-\frac{\sum_{p=1}^{voc}\sum_{n=1}^{doc} x_{pi} log(\lambda_{pi})}{x_{..}}) && (26)
\end{aligned}
$$

4 Pseudo code

Algorithm 1 MCMC Inference for PFA

Require: Train_data,Test_data
Ensure:
Randomly initial all latent variable according to the generative process
Initialize $x_{.ik}, x_{p.k}, x_{..k}, x_{.k.}$ When assigning x_{pi} into each topic.By Eq.(6)
for iter **do**
  for $topic\_index = 1 \to K$ **do**
    Sample ϕ_k by Eq.(12)
    Sample $_k$ by Eq.(17)
    sample l_{ik} CRT by Eq.(20)
    Sample r_k by Eq.(21)
    Sample θ_{ki} by Eq.(24)
  Compute perplexity by Eq.(26)

References

[1] Mingyuan Zhou, Lauren Hannah, David Dunson, and Lawrence Carin. Beta-negative binomial process and poisson factor analysis. In *Artificial Intelligence and Statistics*, pages 1462–1471, 2012.

[2] Mingyuan Zhou and Lawrence Carin. Negative binomial process count and mixture modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(2):307–320, 2013.