ROBUST STREAMING TENSOR FACTORISATION VIA ONLINE VARIATIONAL BAYESIAN INFERENCE

A PREPRINT

October 25, 2019

ABSTRACT

1 Details about Notations

Given data $\mathcal{Y}_{\Omega_T}$ in the form of a tensor of N dimensions, we wish to find its representation in it's CP Factorisation form. $\mathcal{Y}_{\Omega_T} =< A^{(1)}, A^{(2)}, ..., A^{(N)}; b_T > +S_{\Omega_T}$,where $A^{(1)}, .., A^{(N)}$ are the factor matrices which evolve over time ,b_t is temporal factor which follows the first order autoregressive model $p(b_t|J, b_{t-1}) = \mathcal{N}(b_t|Jb_{t-1}, I_r)$ and S_{Ω_T} accounts for the sparse outliers.

2 Probabilistic Model Description

2.1 Likelihood

$$p(Y_{\Omega_T}, \widetilde{D}_{t,\Omega_t})|[A^{(1)}, A^{(2)}, A^{(3)}......A^{(n)}], B, S_{\Omega_T}, \beta) = \prod_{i\in\Omega_t} \mathcal{N}(y_{i_1,i_2,i_3.....i_N}| < \hat{a_{i_1}}^{(1)}, \hat{a_{i_2}}^{(2)}......\hat{a_{i_n}}^{(n)}, b_T > +S_{i_1,i_2.......i_N}, \beta^{-1}) \tag{1}$$

.

$$\widetilde{D}_t = Y_t - S_t$$

2.2 Prior Distributions

$$\prod_{i=1}^{N} p(A^{(n)}|\lambda) = \prod_{k=1}^{N}\prod_{i_k}^{I_k} \mathcal{N}(\hat{a_{i_k}}^{(k)}|0, \Lambda^{-1}) \tag{2}$$

$$\Lambda = \text{Diag}(\lambda)$$

$$p(B|J) = N(b_i; \mu_1, \lambda_1) \prod_{t=2}^{T} \mathcal{N}(b_t; Jb_{t-1}, I_r) \tag{3}$$

$$p(J|\nu) = \prod_{i=1}^{r}\prod_{j=1}^{r} \mathcal{N}(J_{ij}|0, \nu_i^{-1} I) \tag{4}$$

$$p(S_{\Omega_T}|\gamma) = \prod_{i\in\Omega_T} \mathcal{N}(S_{i_1 i_2 ... i_N}|0, \gamma_{i_1 i_2 .... i_N}) \tag{5}$$

$$p(\lambda) = \prod_{i=1}^{r} \frac{1}{\lambda_i} \tag{6}$$

$$p(\nu) = \prod_{i=1}^{r} \frac{1}{\nu_i} \tag{7}$$

$$p(\gamma) = \prod_{i \in \Omega_T} \frac{1}{\gamma_{i_1 i_2 \ldots i_N}} \tag{8}$$

$$p(\beta) = \frac{1}{\beta} \tag{9}$$

3 Updates

3.1 Factor Matrix Updates

$$q(A^{(n)}) = \prod_{i_n=1}^{I_n} \mathcal{N}(\hat{a}_{i_n}^{(n)} | \bar{a}_{i_n}^{(n)}, V_{i_n}^{(n)}) \tag{10}$$

$$V_{i_n}^{(n)} = (\mathop{\mathbb{E}}_q[\beta] \Sigma_{t=i}^{T} \mathop{\mathbb{E}}_q[A_{i_n}^{(\backslash n)T} A_{i_n}^{(\backslash n)T}]_{\Omega t} + \mathop{\mathbb{E}}_q[\Lambda])^{-1} \tag{11}$$

$$\bar{a}_{i_n}^{(n)} = V_{i_n}^{(n)} (\mathop{\mathbb{E}}_q[A_{i_n}^{(\backslash n)T}]_{\Omega_T} \mathrm{vec}(y_{\Omega_T} - \mathop{\mathbb{E}}_q[S_{\Omega_T}])) + \Sigma_{t=i}^{T-1} \mathbb{E}_q[A_{i_n}^{(\backslash n)T}]_{\Omega_T} \mathrm{vec}(\widetilde{D}_{t,\Omega_t}) \tag{12}$$

$$\mathop{\mathbb{E}}_q[A_{i_n}^{(\backslash n)}] = (\mathop{\mathbb{E}}_q[\bigodot_{j \neq n} A^{(j)}])_{I_{i_n}} \tag{13}$$

3.2 Updates for J

$$\Xi^J = (\mathrm{Diag}(\hat{\nu}) + \sum_{t=1}^{T-1} \Sigma_{t,t-1}^B)^{-1} \tag{14}$$

$$\mu_i^J = [\Xi_J \vec{\Sigma}_{t,t-1}^B]_{:,i} \tag{15}$$

$$\Sigma_{t,t-1}^B = \mu_t^B (\mu_{t-1}^B)^T + \Xi_{t,t-1}^B \tag{16}$$

3.3 Updates for B

$$\hat{\beta} = \frac{|\Omega|}{\sum_{t=1}^{T} \sum_{i \in \Omega_t} y_{i\tau}^2 - 2(y_{i\tau} - \mu_e^{i\tau})(\mu_i^A)^T \mu_\tau^B - 2 y_{i\tau} \mu_e^{i\tau} + (\mu_e^{i\tau})^2 + \Xi \Xi_e^{i\tau} + \mathrm{Trace}(\Sigma_i^A \Sigma_{\tau,\tau}^B)} \tag{17}$$

$$\Xi_B = \hat{\beta} \cdot \mathrm{Diag}(\Xi_1^A, \Xi_2^A \ldots \Xi_\tau^A) + \begin{bmatrix} \Lambda^{-1} & -\hat{J} & \ldots & 0 \\ -\hat{J} & \ldots. & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{bmatrix} \tag{18}$$

$$\mu_B = \Xi_B \begin{bmatrix} \hat{\beta} \sum_{i \in \Omega_1} y_{i1} \mu_i^A + \Lambda_1^{-1} \mu_1 \\ \hat{\beta} \sum_{i \in \Omega_2} \\ \ldots \\ \ldots \\ \ldots \\ \hat{\beta} \sum_{i \in \Omega_t} y_{it} \mu_i^A \end{bmatrix} \tag{19}$$

3.4 Updates for S

$$q(S_{\Omega_T}) = \prod_{i_1 i_2..i_N \in \Omega_T} \mathcal{N}(S_{i_1 i_2..i_N} | \bar{S}_{i_1 i_2..i_N, \sigma^2_{i_1 i_2..i_N}}) \tag{20}$$

$$\bar{S}_{i_1 i_2..i_N} = \sigma^2_{i_1 i_2..i_N} \mathbb{E}_q[\beta](\mathcal{Y}_{i_1 i_2..i_N} - \mathbb{E}_q[< \hat{a}^{(1)}_{i_1}, ..., \hat{a}^{(N)}_{i_N}; b_T >]) \tag{21}$$

$$\sigma^2_{i_1 i_2..i_N} = (\mathop{\mathbb{E}}_q[\gamma_{i_1 i_2..i_N}] + \mathop{\mathbb{E}}_q[\beta)^{-1} \tag{22}$$

3.5 Update for γ

$$\gamma_{i_1,i_2,i_3.....i_N} = \Gamma(\gamma_{i_1,i_2,i_3...i_N} | \frac{1}{2}, \frac{\mathbb{E}_q(S^2_{i_1,i_2......i_N})}{2}) \tag{23}$$

$$< \gamma_{i_1,i_2,....i_N} >= \frac{1}{S^2_{i_1,i_2......i_N} + \sigma^2_{i_1,i_2....i_N}} \tag{24}$$

3.6 Update for β

$$q^\beta = \Gamma(\beta | a_0, b_0) \tag{25}$$

$$a_0 = \frac{|\Omega_T|}{2} \tag{26}$$

$$b_0 = \frac{\mathbb{E}_q[||\mathcal{Y} - < \hat{a}^{(1)}_{i_1}, \hat{a}^{(2)}_{i_2}, ..., \hat{a}^{(N)}_{i_N}; b_T > - S_T||^2_F]}{2} \tag{27}$$

3.7 Update for ν

$$\hat{\nu} = \frac{m}{\sum_{k \in \text{cols}} [\mu^J_k]^2 + [\Xi^J_k]_{ii}} \tag{28}$$