"Tensor RPCA by Bayesian CP Factorization with Complex Noise": Supplementary Material

Qiong Luo[1,2], Zhi Han[1]*, Xi'ai Chen[1,2], Yao Wang[3], Deyu Meng[3], Dong Liang[3], Yandong Tang[1]
[1]State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences;
[2]University of Chinese Academy of Sciences; [3]Xi'an Jiaotong University
{luoqiong, hanzhi, chenxiai, ytang}@sia.cn, {dymeng, liangdong}@mail.xjtu.edu.cn, yao.s.wang@gmail.com

Abstract

The supplementary material includes the TenRPCA-MoG full hierarchical Bayesian model and the detail of variational inference process for inferring the posterior of the model.

1. Hierarchical model for TenRPCA-MoG

The TenRPCA-MoG

$$\mathcal{Y} = \mathcal{X} + \mathcal{E}$$

where y_{ijk} is the elements of tensor $\mathcal{Y}$ in the i-th row, j-th column and k-th tube. Based on CP decomposition, the elementary of the three order tensor $\mathcal{X}$ can be expressed by:

$$x_{ijk} = \sum_{d=1}^{r} u_{id} v_{jd} t_{kd}$$

where u_{id} is i-th row and d-th column of factor matrix U, as well as v_{jd} and t_{kd}. For inference covenience, it is often written as the following form:

$$x_{ijk} = u_{id}(v_{jd} \odot t_{kd})^T$$

We formulate the full hierarchical form of the TenRPCA-MoG as:

$$y_{ijk} = u_{id}(v_{jd} \odot t_{kd})^T + e_{ijk}$$

$$\mathbf{u}_{.d} \sim \mathcal{N}(\mathbf{u}_{.d}|0, \gamma_d^{-1}\mathbf{I}_f)$$

$$\mathbf{v}_{.d} \sim \mathcal{N}(\mathbf{v}_{.d}|0, \gamma_d^{-1}\mathbf{I}_g)$$

$$\mathbf{t}_{.d} \sim \mathcal{N}(\mathbf{t}_{.d}|0, \gamma_d^{-1}\mathbf{I}_m)$$

$$\gamma_d \sim Gam(\gamma_d|a_0, b_0)$$

$$e_{ijk} \sim \prod_{n=1}^{N} \mathcal{N}(e_{ijk}|\mu_n, {\tau_n}^{-1})^{z_{ijkn}}$$

$$\mathbf{z}_{ijk} \sim Multinomia(\mathbf{z}_{ijk}|\boldsymbol{\pi})$$

$$\boldsymbol{\pi} \sim Dir(\boldsymbol{\pi}|\boldsymbol{\alpha}_0)$$

$$\mu_n, \tau_n \sim \mathcal{N}(\mu_n|\mu_0, (\beta_0\tau_n)^{-1}) Gam(\tau_n|c_0, d_0)$$

The full likelihood of the generative model can be formulated by:

$$\begin{aligned}
&p(U, V, T, \mathcal{Z}, \mu, \tau, \pi, \gamma, \mathcal{Y}) \\
&= p(\mathcal{Y}|U, V, T, \mathcal{Z}, \mu, \tau) p(\mathcal{Z}|\pi) p(\mu|\tau) p(\tau) p(U|\gamma) \\
&\quad p(V|\gamma) p(T|\gamma) p(\gamma) \\
&= \textstyle\prod_{ijk} \prod_{n=1}^{N} p(y_{ijk}|u_{i.}, v_{j.}, t_{k.}, \mu_n, \tau_n^{-1})^{z_{ijkn}} \\
&\quad \textstyle\prod_{ijk} p(z_{ijk}|\pi) P(\pi) \prod_{n=1}^{N} p(\mu_n, \tau_n) \\
&\quad \textstyle\prod_{d=1}^{R} \{p(u_{.d}|\gamma_d) p(v_{.d}|\gamma_d) p(t_{.d}|\gamma_d) p(\gamma_d)\} \\
&= \textstyle\prod_{ijk} \prod_{n=1}^{N} \mathcal{N}\big(y_{ijk}|u_{id}(v_{jd} \odot t_{kd})^T + \mu_n, \tau_n^{-1}\big)^{z_{ijk}} \\
&\quad \textstyle\prod_{ijk} \prod_{n=1}^{N} \pi^{z_{ijk}} Dir(\pi|\alpha_0) \\
&\quad \textstyle\prod_{n=1}^{N} \Big\{\mathcal{N}(\mu_n|\mu_0, (\beta_0\tau_n)^{-1}) Gam(\tau_n|c_0, d_0)\Big\} \\
&\quad \textstyle\prod_{d=1}^{D} \{\mathcal{N}(u_{.d}|0, {\gamma_d}^{-1} I_f) \mathcal{N}(v_{.d}|0, {\gamma_d}^{-1} I_g) \\
&\quad \mathcal{N}(t_{.d}|0, {\gamma_d}^{-1} I_m) Gam(\gamma_d|a_0, b_0)
\end{aligned}$$

2. Update equations

The variational update equations for the posterior of all the involved parameters in the proposed model has the following closed form.
Infer U:

$$q(\mathbf{u}_{i.}) = \mathcal{N}(\mathbf{u}_{i.}|\mu_{\mathbf{u}_{i.}}, \textstyle\sum_{\mathbf{u}_{i.}})$$

where $\langle . \rangle$ denotes the expectation, and

$$\textstyle\sum_{\mathbf{u}_{i.}} = \Big\{\sum_n \langle\tau_n\rangle \sum_{jk} \langle z_{ijkn}\rangle \big\langle \big(\mathbf{t}_{k.}{}^T \mathbf{t}_{k.}\big) * \big(\mathbf{v}_{j.}{}^T \mathbf{v}_{j.}\big)\big\rangle + \mathbf{\Gamma}\Big\}^{-1}$$

$$\begin{aligned}
\mu_{\mathbf{u}_{i.}}^T &= \textstyle\sum_{\mathbf{u}_{i.}} * \Big\{\sum_n \langle\tau_n\rangle \sum_{jk} \langle z_{ijkn}\rangle (y_{ijk} - \langle\mu_n\rangle) \\
&\quad \langle \mathbf{t}_{k.} \odot \mathbf{v}_{j.}\rangle\Big\}^T
\end{aligned}$$

Infer V:

$$q(\mathbf{v}_{j.}) = \mathcal{N}(\mathbf{v}_{j.}|\mu_{\mathbf{v}_{j.}}, \textstyle\sum_{\mathbf{v}_{j.}})$$

*Corresponding author.

where

$$\sum_{\mathbf{v}_{j.}} = \left\{\sum_n \langle\tau_n\rangle \sum_{ik} \langle z_{ijkn}\rangle \left\langle \left(\mathbf{t}_{k.}{}^T\mathbf{t}_{k.}\right) * \left(\mathbf{u}_{i.}{}^T\mathbf{u}_{i.}\right)\right\rangle + \mathbf{\Gamma}\right\}^{-1}$$

$$\mu^T_{\mathbf{v}_{j.}} = \sum_{\mathbf{v}_{j.}} * \{\sum_n \langle\tau_n\rangle \sum_{ik} \langle z_{ijkn}\rangle (y_{ijk} - \langle\mu_n\rangle) \langle \mathbf{t}_{k.} \odot \mathbf{u}_{i.}\rangle\}^T$$

Infer T:

$$q(\mathbf{t}_{k.}) = \mathcal{N}(\mathbf{t}_{k.}|\mu_{\mathbf{t}_{k.}}, \sum\nolimits_{\mathbf{t}_{k.}})$$

where

$$\sum_{\mathbf{t}_{k.}} = \left\{\sum_n \langle\tau_n\rangle \sum_{ij} \langle z_{ijkn}\rangle \left\langle \left(\mathbf{v}_{j.}{}^T\mathbf{v}_{j.}\right) * \left(\mathbf{u}_{i.}{}^T\mathbf{u}_{i.}\right)\right\rangle + \mathbf{\Gamma}\right\}^{-1}$$

$$\mu^T_{\mathbf{t}_{k.}} = \sum_{\mathbf{t}_{k.}} * \left\{\sum_n \langle\tau_n\rangle \sum_{ij} \langle z_{ijkn}\rangle (y_{ijk} - \langle\mu_n\rangle) \langle \mathbf{v}_{j.} \odot \mathbf{u}_{i.}\rangle\right\}^T$$

Infer γ

$$q(\gamma_d) = Gam(\gamma_d|a_d, b_d)$$

where

$$a_d = a_0 + \frac{f+g+m}{2}$$

$$b_d = b_0 + \frac{1}{2}\left(\left\langle \mathbf{u}^T_{.d}\mathbf{u}_{.d}\right\rangle + \left\langle \mathbf{v}^T_{.d}\mathbf{v}_{.d}\right\rangle + \left\langle \mathbf{t}^T_{.d}\mathbf{t}_{.d}\right\rangle\right)$$

Infer $\mathcal{Z}$

$$q(\boldsymbol{z}_{ijk}) = \prod_n r_{ijkn}{}^{z_{jikn}}$$

where

$$r_{ijkn} = \frac{\rho_{ijkn}}{\sum_n \rho_{ijkn}}$$

$$\ln\rho_{jikn} = \tfrac{1}{2}\langle\ln\tau_n\rangle - \tfrac{1}{2}\ln 2\pi - \tfrac{1}{2}\langle\tau_n\rangle \left\langle (y_{jik} - \mathbf{u}_{i.}\left(\mathbf{t}_{k.} \odot \mathbf{v}_{j.}\right)^T - \mu_n\Big)\right\rangle + \langle\ln\pi_n\rangle$$

Infer μ, τ

$$q(\mu_n, \tau_n) = \mathcal{N}(m_n, (\beta_n\tau_n)^{-1}) Gam(\tau_n|c_n, d_n)$$

where

$$\beta_n = \beta_0 + \sum_{ijk} \langle z_{ijkn}\rangle$$

$$m_n = \frac{1}{\beta_n}(\beta_0\mu_0 + \sum_{ijk} \langle z_{ijkn}\rangle (y_{ijk} - \langle u_{i.}\rangle \langle v_{j.} \odot t_{k.}\rangle^T)$$

$$c_n = c_0 + \frac{1}{2}\sum_{ijk} \langle z_{ijkn}\rangle$$

$$d_n = d_0 + \tfrac{1}{2}\left\{\sum_{ijk} \langle z_{ijkn}\rangle \left\langle \left(y_{ijk} - \langle u_{i.}\rangle \langle v_{j.} \odot t_{k.}\rangle^T\right)^2\right\rangle + \beta_0\mu_0{}^2 - \tfrac{1}{\beta_n}\left(\sum_{ijk} \langle z_{ijkn}\rangle (y_{ijk} - \langle u_{i.}\rangle \langle v_{j.} \odot t_{k.}\rangle^T) + \beta_0\mu_0\right)^2\right\}$$

Infer π

$$q(\boldsymbol{\pi}) = Dir(\boldsymbol{\pi}|\boldsymbol{\alpha})$$

where

$$\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n)$$

$$\alpha_n = \alpha_{0n} + \sum_{ijk} \langle z_{ijkn}\rangle$$

3. Calculation of expectations

The calculation of the expectations for the variational update equations under the given variational distributions has the following closed form:

$$\langle\tau_n\rangle = \tfrac{c_n}{d_n}$$
$$\langle z_{ijkn}\rangle = r_{ijkn}$$
$$\langle\ln\tau_n\rangle = \psi(c_n) - lnd_n$$
$$\langle\ln\pi_n\rangle = \psi(\alpha_n) - \psi(\hat{\alpha}_n), \hat{\alpha} = \sum_k^N \alpha_n$$
$$\left\langle \left(y_{ijk} - \mathbf{u}_{i.}(\mathbf{v}_{j.} \odot \mathbf{t}_{k.})^T - \mu_n\right)^2\right\rangle = \left\langle \left(y_{ijk} - \mathbf{u}_{i.}(\mathbf{v}_{j.} \odot \mathbf{t}_{k.})^T\right)^2\right\rangle - 2\langle\mu_n\rangle \left(y_{ijk} - \langle\mathbf{u}_{i.}\rangle \langle\mathbf{v}_{j.} \odot \mathbf{t}_{k.}\rangle^T\right) + \left\langle\mu_n{}^2\right\rangle$$
$$\langle\mu_n\rangle = m_n$$
$$\left\langle\mu_n{}^2\right\rangle = (\beta_n\tau_n)^{-1} + m_n{}^2$$
$$\left\langle \left(y_{ijk} - \mathbf{u}_{i.}(\mathbf{v}_{j.} \odot \mathbf{t}_{k.})^T\right)^2\right\rangle = y_{ijk}{}^2 + tr\left(\left\langle\mathbf{u}_{i.}{}^T\ \mathbf{u}_{i.}\right\rangle\ \left\langle\mathbf{v}_{j.}{}^T\ \mathbf{v}_{j.}\right\rangle \left\langle\mathbf{t}_{k.}{}^T\mathbf{t}_{\ k.}\right\rangle\right) - 2y_{ijk}\langle\mathbf{u}_{i.}\rangle \langle\mathbf{v}_{j.} \odot \mathbf{t}_{k.}\rangle^T$$
$$\left\langle\mathbf{u}_{i.}{}^T\ \mathbf{u}_{i.}\right\rangle = \sum\nolimits_{\mathbf{u}_{i.}} + \langle\mathbf{u}_{i.}\rangle \langle\mathbf{u}_{i.}\rangle^T$$
$$\left\langle\mathbf{v}_{j.}{}^T\ \mathbf{v}_{j.}\right\rangle = \sum\nolimits_{\mathbf{v}_{j.}} + \langle\mathbf{v}_{j.}\rangle \langle\mathbf{v}_{j.}\rangle^T$$
$$\left\langle\mathbf{t}_{k.}{}^T\ \mathbf{t}_{k.}\right\rangle = \sum\nolimits_{\mathbf{t}_{k.}} + \langle\mathbf{t}_{k.}\rangle \langle\mathbf{t}_{k.}\rangle^T$$
$$\mathbf{\Gamma} = diag\left(\langle\gamma\rangle\right), \langle\gamma_d\rangle = \tfrac{a_d}{b_d}$$
$$\left\langle\mathbf{u}_{.d}{}^T\ \mathbf{u}_{.d}\right\rangle = \langle\mathbf{u}_{.d}\rangle^T \langle\mathbf{u}_{.d}\rangle + \sum_{i=1}^f \left(\sum\nolimits_{\mathbf{u}_{.d}}\right)_{dd}$$
$$\left\langle\mathbf{v}_{.d}{}^T\ \mathbf{v}_{.d}\right\rangle = \langle\mathbf{v}_{.d}\rangle^T \langle\mathbf{v}_{.d}\rangle + \sum_{j=1}^g \left(\sum\nolimits_{\mathbf{v}_{.d}}\right)_{dd}$$
$$\left\langle\mathbf{t}_{.d}{}^T\ \mathbf{t}_{.d}\right\rangle = \langle\mathbf{t}_{.d}\rangle^T \langle\mathbf{t}_{.d}\rangle + \sum_{k=1}^m \left(\sum\nolimits_{\mathbf{t}_{.d}}\right)_{dd}$$

where $\psi(.)$ denotes the digamma function defined by $\psi(x) = \frac{d}{dx}\ln\Gamma(x)$.