Article

Hyperspectral Image Restoration Under Complex Multi-band Noises

Zongsheng Yue [1], Deyu Meng[1], Yongqing Sun[2] and Qian Zhao [1]*

[1] Institute for Information and System Sciences and Ministry of Education Key Lab of Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an, Shaan'xi, 710049 P.R. China; zsy20110806207@stu.xjtu.edu.cn; dymeng@xjtu.edu.cn

[2] NTT Media Intelligence Labs, Japan; yongqing.sun@lab.ntt.co.jp

* Correspondence: timmy.zhaoqian@gmail.com; Tel.: +86-138-9187-3702

Academic Editor: name

Version August 31, 2018 submitted to Remote Sens.

Abstract: In this supplementary material, we provide more details on the computations involved in the proposed variational inference algorithm.

1. Variational Assumation

The full likelihood of the proposed NIID-MSL model is expressed as:

$$
\begin{aligned}
& p(\boldsymbol{U},\boldsymbol{V},\boldsymbol{\xi},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\alpha},\boldsymbol{\lambda},\gamma,\boldsymbol{Y})\\
&= p(\boldsymbol{Y}|\boldsymbol{U},\boldsymbol{V},\boldsymbol{\xi},\boldsymbol{C},\boldsymbol{Z})p(\boldsymbol{U}|\boldsymbol{\lambda})p(\boldsymbol{V}|\boldsymbol{\lambda})p(\boldsymbol{\lambda})p(\boldsymbol{\xi})p(\boldsymbol{C}|\boldsymbol{\beta})p(\boldsymbol{\beta}'|\gamma)p(\gamma)p(\boldsymbol{Z}|\boldsymbol{\pi})p(\boldsymbol{\pi}'|\boldsymbol{\alpha})p(\boldsymbol{\alpha})\\
&= \prod_{i,j,t,k}\left\{\mathcal{N}\left(y_{ij}\middle|\boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T+e_{ij},\xi_k^{-1}\right)^{\mathbf{1}[c_{jt}=k]}\right\}^{\mathbf{1}[z_{ij}=t]}\prod_{i,j}Multi(z_{ij}|\boldsymbol{\pi}_j)\prod_{j,t}Multi(c_{jt}|\boldsymbol{\beta})\prod_k Gam(\xi_k|a_0,b_0)\\
&\prod_r\mathcal{N}\left(\boldsymbol{u}_{\cdot r}\middle|\mathbf{0},{\lambda_r}^{-1}\boldsymbol{I}_d\right)\mathcal{N}\left(\boldsymbol{v}_{\cdot r}\middle|\mathbf{0},{\lambda_r}^{-1}\boldsymbol{I}_{\boldsymbol{B}}\right)Gam(\lambda_r^v|p_0,q_0)\prod_{j,t}Beta({\pi_{jt}}'|1,\alpha_j)\prod_j Gam(\alpha_j|e_0,f_0)\\
&\prod_k Beta(\beta_k'|1,\gamma)Gam(\gamma|c_0,d_0).
\end{aligned}
$$

In the maintext, we have introduced the variational inference to calculate posterior of this model and assumed the approximation of posterior have a factorized form as follows:

$$
\begin{aligned}
& q(\boldsymbol{U},\boldsymbol{V},\boldsymbol{\xi},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\lambda},\boldsymbol{\alpha},\gamma)\\
&= \prod_{i=1}^{d}q(\boldsymbol{u}_{i\cdot})\prod_{j=1}^{B}q(\boldsymbol{v}_{j\cdot})\prod_{k=1}^{K}q(\xi_k)q(\beta_k')\prod_{i,j}^{d,B}q(z_{ij})\prod_{j=1}^{B}\prod_{t=1}^{T}q(c_{jt})q(\pi_{jt}')\prod_{r=1}^{l}q(\lambda_r)\prod_{j=1}^{B}q(\alpha_j)q(\gamma). \qquad (1)
\end{aligned}
$$

2. Update Equations

Next, we give detailed deduction of each factorized distribution involved in posterior of Eq. (1). $E_{\setminus x_i}[f(\boldsymbol{x})]$ denotes the expectation of $f(\boldsymbol{x})$ on set of $\boldsymbol{x}$ with x_i removed.

Infer C and Z:

$$\begin{aligned}
&\ln q(z_{ij}) \\
&= E_{\backslash \boldsymbol{Z}}\big[p(\boldsymbol{U},\boldsymbol{V},\boldsymbol{\xi},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\alpha},\boldsymbol{\lambda},\gamma,\boldsymbol{Y})\big] + const \\
&= \sum_t \rho_{ijt}\Big\{\sum_k \varphi_{jtk} E_{\backslash \boldsymbol{Z}}\Big[\ln \mathcal{N}\Big(y_{ij} - \boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T\Big|0,\xi_k^{-1}\Big)\Big] + E\big[\ln \pi_{jt}\big]\Big\} + const \\
&= \sum_t \rho_{ijt}\Big\{\sum_k \varphi_{jtk}\Big(-\frac{1}{2}\ln 2\pi + \frac{1}{2}E\big[\ln \xi_k\big] - \frac{1}{2}E\big[\xi_k\big]E\Big[\Big(y_{ij} - \boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T\Big)^2\Big]\Big) + E\big[\ln \pi_{jt}\big]\Big\} + const,
\end{aligned} \tag{2}$$

$$\begin{aligned}
&\ln q(c_{jt}) \\
&= E_{\backslash \boldsymbol{C}}\big[p(\boldsymbol{U},\boldsymbol{V},\boldsymbol{\xi},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\alpha},\boldsymbol{\lambda},\gamma,\boldsymbol{Y})\big] + const \\
&= \sum_k \varphi_{jtk}\Big\{\sum_i \rho_{ijt}^v E_{\backslash \boldsymbol{C}}\Big[\ln \mathcal{N}\Big(y_{ij} - \boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T\Big|0,\xi_k^{-1}\Big)\Big] + E\big[\ln \beta_k\big]\Big\} + const \\
&= \sum_k \varphi_{jtk}\Big\{\sum_i \rho_{ijt}\Big(-\frac{1}{2}\ln 2\pi + \frac{1}{2}E\big[\ln \xi_k\big] - \frac{1}{2}E\big[\xi_k\big]E\Big[\Big(y_{ij} - \boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T\Big)^2\Big]\Big) + E\big[\ln \beta_k\big]\Big\} + const,
\end{aligned} \tag{3}$$

Taking the exponential of both sides of Eq. (2), Eq. (3) and normalizing the right side, we obtain

$$q(z_{ij}) = Multi(\boldsymbol{\rho_{ij}}), \quad q(c_{jt}) = Multi(\boldsymbol{\varphi_{jt}}),$$

where

$$\begin{aligned}
\varphi_{jtk} &\propto \exp\left\{\sum_{i=1}^d \rho_{ijt}\left(\frac{1}{2}E[\ln \xi_k] - \frac{1}{2}\ln 2\pi - \frac{1}{2}E[\xi_k]E[(y_{ij} - \boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T)^2]\right) + E[\ln \beta_k]\right\}, \\
\rho_{ijt} &\propto \exp\left\{\sum_{k=1}^K \varphi_{itk}\left(\frac{1}{2}E[\ln \xi_k] - \frac{1}{2}\ln 2\pi - \frac{1}{2}E[\xi_k]E[(y_{ij} - \boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T)^2]\right) + E[\ln \pi_{jt}]\right\}, \\
E[\ln \beta_k] &= E[\ln \beta_k'] + \sum_{l=1}^{k-1} E[\ln(1-\beta_l')], \;\; E[\ln \pi_{jt}] = E[\ln \pi_{jt}'] + \sum_{s=1}^{t-1} E[\ln(1-\pi_{js}')].
\end{aligned}$$

Infer ξ:

$$\begin{aligned}
&\ln q(\xi_k) \\
&= E_{\backslash \boldsymbol{\xi}}\big[p(\boldsymbol{U},\boldsymbol{V},\boldsymbol{\xi},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\alpha},\boldsymbol{\lambda},\gamma,\boldsymbol{Y})\big] + const \\
&= \sum_{i,j,t} \rho_{ijt}\varphi_{jtk} E_{\backslash \boldsymbol{\xi_k}}\Big[\ln \mathcal{N}\Big(x_{ij} - \boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T\Big|0,\xi_k^{-1}\Big)\Big] + (a_0 - 1)\ln \xi_k - b_0\xi_k \\
&= \Big(\frac{1}{2}\sum_{i,j,t}\rho_{ijt}\varphi_{jtk} + a_0 - 1\Big)\ln \xi_k - \Big\{\frac{1}{2}\sum_{i,j,t}\rho_{ijt}\varphi_{jtk} E\Big[\Big(x_{ij} - \boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T\Big)^2\Big] + b_0\Big\}\xi_k + const,
\end{aligned} \tag{4}$$

Aftering taking exponential of both side of Eq. (4), we have:

$$q(\xi_k) = Gam(\xi_k | a_k, b_k),$$

where

$$a_k = \frac{1}{2}\sum_{i,j,t}\rho_{ijt}\varphi_{jtk} + a_0,$$
$$b_k = \frac{1}{2}\sum_{i,j,t}\rho_{ijt}\varphi_{jtk}E\left[\left(y_{ij} - \boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T\right)^2\right] + b_0.$$

Infer $\boldsymbol{\pi}'$ and $\boldsymbol{\beta}'$:

$$\begin{aligned}
&\ln q(\pi_{jt}{}') \\
&= E_{\setminus\boldsymbol{\pi}}\left[p(\boldsymbol{U},\boldsymbol{V},\boldsymbol{\xi},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\alpha},\boldsymbol{\lambda},\gamma,\boldsymbol{Y})\right] + const \\
&= \sum_i \rho_{ijt}\ln\pi'_{jt} + (\alpha_j - 1)\ln(1-\pi'_{jt}) + const \\
&= \left(\sum_{i,s=t+1}\rho_{ijs} + \alpha_j - 1\right)\ln(1-\pi'_{jt}) + \left(\sum_i \rho_{ijt}\right)\ln\pi'_{jt} + const,
\end{aligned} \tag{5}$$

then we take exponential of both side of Eq. (5) and can get:

$$q(\pi'_{jt}) = Beta(\pi'_{jt}|r^1_{jt}, r^2_{jt}),$$

where

$$r^1_{jt} = \sum_i \rho_{ijt} + 1,\ \ r^2_{jt} = \sum_{i,s=t+1}\rho_{ijs} + \alpha_j.$$

Similarly, we have:

$$q(\beta'_k) = Beta(\beta'_k|s^1_k, s^2_k),$$

where

$$s^1_k = \sum_{j,t}\varphi_{jtk} + 1,\ \ s^2_k = \sum_{j,t,l=k+1}\varphi_{jtl} + \gamma.$$

Infer $\boldsymbol{\alpha}$ and γ:

$$\begin{aligned}
&\ln q(\alpha_j) \\
&= E_{\setminus\boldsymbol{\alpha}}\left[p(\boldsymbol{U},\boldsymbol{V},\boldsymbol{\xi},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\alpha},\boldsymbol{\lambda},\gamma,\boldsymbol{Y})\right] + const \\
&= \sum_t \left((\alpha_j - 1)E[\ln(1-\pi'_{jt})] + \ln\alpha_j\right) + (e_0 - 1)\ln\alpha_j - f_0\alpha_j + const \\
&= (T + e_0 - 1)\ln\alpha_j - \left(f_0 - \sum_t E\left[\ln(1-\pi'_{jt})\right]\right) + const,
\end{aligned} \tag{6}$$

From Eq. (6), We can easily get the following equations of $\boldsymbol{\alpha}$:

$$q(\alpha_j) = Gam(\alpha_j|e_j, f_j),$$

where

$$e_j = T + e_0,\ \ f_j = f_0 - \sum_t E\left[\ln(1-\pi'_{jt})\right].$$

Similarly, we can update variable γ as follows:

$$q(\gamma) = Gam(\gamma|c,d),$$

where

$$c = K + c_0, \;\; d = d_0 - \sum_k E\left[\ln(1-\beta_k^{'})\right].$$

Infer $\boldsymbol{U}$ and $\boldsymbol{V}$:

$$\begin{aligned}
&\ln q(\boldsymbol{u}_{i\cdot}) \\
&= E_{\backslash \boldsymbol{U}}\left[p(\boldsymbol{L},\boldsymbol{V},\boldsymbol{\xi},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\alpha},\boldsymbol{\lambda},\gamma,\boldsymbol{Y})\right] + const \\
&= \sum_{j,t,k} \rho_{ijt}\varphi_{jtk}\left[-\frac{1}{2}E[\xi_k]E\left[\left(y_{ij} - \boldsymbol{u}_{i\cdot}\boldsymbol{v}_{j\cdot}^T\right)^2\right]\right] - \frac{1}{2}\boldsymbol{u}_{i\cdot}\boldsymbol{\Lambda}\boldsymbol{u}_{i\cdot}^T + const,
\end{aligned} \tag{7}$$

where $\boldsymbol{\Lambda} = diag(E[\boldsymbol{\lambda}])$. Taking exponential of both sides of Eq. (7), and normalizing the result, we obtain the posterior distribution of $\boldsymbol{u}_{i\cdot}$:

$$q(\boldsymbol{u}_{i\cdot}) = \mathcal{N}(\boldsymbol{u}_{i\cdot}|\boldsymbol{\mu}_{i\cdot}^u, \boldsymbol{\Sigma}_i^u),$$

where

$$\boldsymbol{\Sigma}_i^u = \left(\sum_{j,t,k}\rho_{ijt}\varphi_{jtk}E[\xi_k]E[\boldsymbol{v}_{j\cdot}^T\boldsymbol{v}_{j\cdot}] + \boldsymbol{\Lambda}\right)^{-1}, \;\; \boldsymbol{\mu}_{i\cdot}^u = \sum_{j,t,k}\rho_{ijt}\varphi_{jtk}E[\xi_k]y_{ij}\boldsymbol{v}_{j\cdot}\boldsymbol{\Sigma}_i^u.$$

Similarly, each column of $\boldsymbol{V}$ is also a Gaussian distribution, i.e.,

$$q(\boldsymbol{v}_{j\cdot}) = \mathcal{N}(\boldsymbol{v}_{j\cdot}|\boldsymbol{\mu}_{j\cdot}^v, \boldsymbol{\Sigma}_j^v),$$

where

$$\boldsymbol{\Sigma}_j^v = \left(\sum_{i,t,k}\rho_{ijt}\varphi_{jtk}E[\xi_k]E[\boldsymbol{u}_{i\cdot}^T\boldsymbol{u}_{i\cdot}] + \boldsymbol{\Lambda}\right)^{-1}, \;\; \boldsymbol{\mu}_{j\cdot}^v = \sum_{i,t,k}\rho_{ijt}\varphi_{jtk}E[\xi_k]y_{ij}\boldsymbol{u}_{i\cdot}\boldsymbol{\Sigma}_j^v.$$

Infer λ:

$$\begin{aligned}
&\ln q(\lambda_r) \\
&= E_{\backslash \lambda}\left[p(\boldsymbol{U},\boldsymbol{V},\boldsymbol{\xi},\boldsymbol{C},\boldsymbol{Z},\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\alpha},\boldsymbol{\lambda},\gamma,\boldsymbol{Y})\right] + const \\
&= \left(\frac{d}{2} + \frac{B}{2} + p_0 - 1\right)\ln\lambda_r - \left(\frac{1}{2}E\left[\boldsymbol{u}_{\cdot r}{}^T\boldsymbol{u}_{\cdot r}\right] + \frac{1}{2}E\left[\boldsymbol{v}_{\cdot r}{}^T\boldsymbol{v}_{\cdot r}\right] + q_0\right)\lambda_r + const,
\end{aligned} \tag{8}$$

Thus, we can get the following Updated equations:

$$q(\lambda_r) = Gam(\lambda_r|p_r, q_r),$$

where

$$p_r = \frac{d}{2} + \frac{B}{2} + p_0, \;\; q_r = \frac{1}{2}E[\boldsymbol{u}_{\cdot r}{}^T\boldsymbol{u}_{\cdot r}] + \frac{1}{2}E[\boldsymbol{v}_{\cdot r}{}^T\boldsymbol{v}_{\cdot r}] + q_0.$$

3. Calculation of Expectations

The expectation in the variational update equations can be calculated with respect to the current variational distributions, as listed in the followings:

$$
\begin{aligned}
E[\xi_k] &= \frac{a_k}{b_k}, \\
E\left[\ln \xi_k\right] &= \psi(a_k) - \ln b_k, \\
E\left[\ln \pi_{jt}^{'}\right] &= \psi\left(r_{jt}^{1}\right) - \psi\left(r_{jt}^{1} + r_{jt}^{2}\right), \\
E\left[\ln(1-\pi_{jt}^{'})\right] &= \psi\left(r_{jt}\right) - \psi\left(r_{jt}^{1} + r_{jt}^{2}\right), \\
E\left[\ln \pi_{jt}\right] &= E\left[\ln \pi_{jt}^{'}\right] + \sum_{s=1}^{t-1} E\left[\ln(1-\pi_{js}^{'})\right], \\
E\left[\ln \beta_{k}^{'}\right] &= \psi\left(s_{k}^{1}\right) - \psi\left(s_{k}^{1} + s_{k}^{2}\right), \\
E\left[\ln\left(1-\beta_{k}^{'}\right)\right] &= \psi\left(s_{k}^{2}\right) - \psi\left(s_{k}^{1} + s_{k}^{2}\right), \\
E\left[\ln \beta_k\right] &= E\left[\ln \beta_{k}^{'}\right] + \sum_{l=1}^{k-1} E\left[\ln(1-\beta_{l}^{'})\right], \\
E\left[\boldsymbol{u}_{i\cdot}^{T}\boldsymbol{u}_{i\cdot}\right] &= \boldsymbol{\mu}_{i\cdot}^{u\,T}\boldsymbol{\mu}_{i\cdot}^{u} + \boldsymbol{\Sigma}_{i}^{u}, \\
E\left[\boldsymbol{v}_{j\cdot}^{T}\boldsymbol{v}_{j\cdot}\right] &= \boldsymbol{\mu}_{j\cdot}^{v\,T}\boldsymbol{\mu}_{j\cdot}^{v} + \boldsymbol{\Sigma}_{j}^{v}, \\
E\left[\boldsymbol{u}_{\cdot r}^{T}\boldsymbol{u}_{\cdot r}\right] &= \boldsymbol{\mu}_{\cdot r}^{u\,T}\boldsymbol{\mu}_{\cdot r}^{u} + \sum_{i}\left(\boldsymbol{\Sigma}_{i}^{u}\right)_{rr}, \\
E\left[\boldsymbol{v}_{\cdot r}^{T}\boldsymbol{v}_{\cdot r}\right] &= \boldsymbol{\mu}_{\cdot r}^{v\,T}\boldsymbol{\mu}_{\cdot r}^{v} + \sum_{j}\left(\boldsymbol{\Sigma}_{j}^{v}\right)_{rr}, \\
E\left[\left(y_{ij} - \boldsymbol{u}_{i\cdot}{}^{T}\boldsymbol{v}_{\cdot j}\right)^{2}\right] &= x_{ij}^{2} - 2x_{ij}\boldsymbol{\mu}_{i\cdot}^{u\,T}\boldsymbol{\mu}_{j\cdot}^{v} + tr\left(E\left[\boldsymbol{u}_{i\cdot}^{T}\boldsymbol{u}_{i\cdot}\right]E\left[\boldsymbol{v}_{\cdot j}^{T}\boldsymbol{v}_{\cdot j}\right]\right).
\end{aligned}
$$

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