Robust Multi-view Subspace Learning with Non-identical and Non-independent Distributed Complex Noises

Zongsheng Yue, Hongwei Yong, Deyu Meng, Member, IEEE, Qian Zhao, Yee Leung, and Lei Zhang, Fellow, IEEE

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

I. NIID-MSL MODEL

A. Model Formulation

Basically, we decompose the observed data into:

\[ X^v = S^v + E^v, \]

where \( E^v = \{ e_{ij}^v \}_{d \times n} \) denotes the residual term (i.e., noise component) and \( S^v \in \mathbb{R}^{d \times n} \) is the expected data located on the latent subspace, \( d \) and \( n \) represent the dimensionality and the number of samples in each view.

Firstly, we model the noise term \( E^v \) as follows:

\[
\begin{align*}
\xi_k &\sim \text{Gam}(\varepsilon_0, f_0), \\
e_{ij}^v &\sim \mathcal{N}(0, (\xi_{ij}^v)^{-1}), \\
c_{ij}^v &\sim \text{Multi}(\beta), \\
z_{ij}^v &\sim \text{Multi}(\pi^v), \\
\beta_k &\sim \text{Beta}(1, \gamma), \\
\pi_{ij}^v &\sim \text{Beta}(1, \alpha^v), \\
\gamma &\sim \text{Gam}(\alpha_0, \beta_0), \\
\alpha^v &\sim \text{Gam}(\eta_0, \zeta_0),
\end{align*}
\]

where \( \alpha \) and \( \gamma \) are the concentration parameters, which mainly affect the number of Gaussian components of the second-level GMM in each view and the first-level GMM for the entire dataset, respectively.

As for the expected data \( S^v \), we embedded each view into a latent space \( R \) with a dictionary \( L^v \) as conventional MSL methods, i.e.,

\[
\begin{align*}
S^v &= \sum_{r=1}^{l} L^v_r R_r, \\
R_r &\sim \mathcal{N}(0, \frac{1}{\tau_r} I_n), \\
L^v_r &\sim \mathcal{N}(0, \frac{1}{\lambda^v_r} I_d), \\
\lambda^v_r &\sim \text{Gam}(a_0, b_0), \\
\tau_r &\sim \text{Gam}(c_0, d_0).
\end{align*}
\]

Combining Eqs. (1) - (3), the goal of our proposed NIID-MSL turns to infer the posteriors of all involved variables:

\[ p(L, R, \xi, C, Z, \beta, \pi, \alpha, \lambda, \tau, \gamma | X), \]

where \( C = \{ c_{ij}^v \} \), \( Z = \{ z_{ij}^v \} \).

B. Variational Assumption

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

\[ p(L, R, \xi, C, Z, \beta, \pi, \alpha, \lambda, \tau, \gamma, X) = p(X | L, R, \xi, C, Z) p(L | \lambda) p(\beta | \tau) p(\alpha | \gamma) p(C | \beta) \]

\[ p(\beta | \gamma) p(\pi | \alpha) p(\xi | \lambda) p(\tau | \gamma). \]

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

\[
\begin{align*}
q(L, R, \xi, C, Z, \beta, \pi, \alpha, \lambda, \tau, \gamma) = & \prod_{v=1}^{V} \prod_{i,j} q(z_{ij}^v | \rho_{ij}^v) q(\xi_k | \varepsilon_k, f_k) \\
& \prod_{v=1}^{V} \prod_{i,j} q(c_{ij}^v | \phi_{ij}^v) q(\pi_{ij}^v | r_{ij}^v, w_{ij}^v) q(\beta_k | s_{ki}, s_{ki}) \\
& \prod_{r=1}^{R} q(\lambda_r^v | \alpha_r^v, \beta_r^v) q(\tau_r | c_r, d_r).\end{align*}
\]

Next, we give detailed deduction of each factorized distribution involved in posterior of Eq. (6). \( E_{x_i} f(x_i) \) denotes the expectation of \( f(x_i) \) on set of \( x_i \) with \( x_i \) removed. For notations convenience, we introduced \( \Theta \) to denote all the parameters that need to be inferred, i.e.,

\[
\Theta = \{ L, R, \xi, C, Z, \beta, \pi, \alpha, \lambda, \tau, \gamma, X \}.
\]
Infer $C$ and $Z$:

$$\ln q(x^v_j) = E_{\Theta \varepsilon Z} [p(L, R, \xi, C, Z, \beta, \pi, \alpha, \lambda, \gamma, X)] + \text{const}$$

$$= \sum_t \sum_k \sum \varphi^v_{ik,t} E_{\Theta \varepsilon Z} \left[ \mathcal{N} \left( x^v_{ij} - L^v_i R_j, 0, \xi_k^{-1} \right) \right] + \ln [\text{const}]$$

$$= \sum_t \sum_k \sum \varphi^v_{ik,t} \left[ -\frac{1}{2} \ln 2\pi + \frac{1}{2} E[\ln \xi_k] - \frac{1}{2} E[\xi_k] E \left( \left( x^v_{ij} - L^v_i R_j \right)^2 \right) + \ln [\text{const}] \right]$$

Taking the exponential of both sides of Eq. (7), Eq. (8) and normalizing the right side, we obtain

$$q(x^v_{ij} | \rho^v_{ij}) = \text{Multi}(\rho^v_{ij}), \quad q(c^v_j | \varphi^v_j) = \text{Multi}(\varphi^v_j),$$

where

$$\rho^v_{ij} = \sum_{\alpha} \rho^v_{ij} \varphi^v_{ij}, \quad \varphi^v_{ik} = \sum_{\alpha} \varphi^v_{ik} \varphi^v_{ts},$$

$$\rho^v_{ij} \propto \exp \left\{ \sum_{\alpha} \varphi^v_{ij} \left( \frac{1}{2} \ln 2\pi + \frac{1}{2} E[\ln \xi_k] - \frac{1}{2} E[\xi_k] E \left( \left( x^v_{ij} - L^v_i R_j \right)^2 \right) \right) + \ln [\text{const}] \right\},$$

$$\varphi^v_{ik} \propto \exp \left\{ \sum_{\alpha} \varphi^v_{ik} \left( \frac{1}{2} \ln 2\pi + \frac{1}{2} E[\ln \xi_k] - \frac{1}{2} E[\xi_k] E \left( \left( x^v_{ij} - L^v_i R_j \right)^2 \right) \right) \right\}. $$

Infer $\pi$ and $\beta$:

$$\ln q(\pi^v_j) = E_{\Theta \varepsilon C} [p(L, R, \xi, C, Z, \beta, \pi, \alpha, \lambda, \gamma, X)] + \text{const}$$

$$= \sum_{i,j} \rho^v_{ij} \ln \pi^v_j + \ln(1 - \pi^v_j) + \text{const}$$

After taking exponential of both side of Eq. (13), we have:

$$q(\pi^v_j | e_k, f_k) = \text{Gam}(\pi^v_j | e_k, f_k),$$

where

$$e_k = \frac{1}{2} \sum_{i,j} \rho^v_{ij} + e_0,$$

$$f_k = \frac{1}{2} \sum_{i,j} \rho^v_{ij} E \left( (x^v_{ij} - L^v_i R_j)^2 \right) + f_0.$$ 

Infer $\alpha$ and $\gamma$:

$$\ln q(\alpha^v) = \sum_{i,j} \left( (\alpha^v - 1) E[\ln(1 - \pi^v_j)] + \ln \alpha^v \right) + (m_0 - 1) \ln \alpha^v$$

$$= (T + m_0 - 1) \ln \alpha^v - \left( n_0 - \sum_t E[\ln(1 - \pi^v_j)] \right) + \text{const},$$

$$= \text{Gam}(\alpha^v | m^v, n^v),$$

where

$$m^v = T + m_0,$$

$$n^v = n_0 - \sum_t E[\ln(1 - \pi^v_j)].$$
Similarly, we can update variable $\gamma$ as follows:

$$q(\gamma | g, h) = \text{Gam}(\gamma | g, h),$$

where

$$g = K + g_0, \quad h = h_0 - \sum_k E \left[ \ln(1 - \beta_k^\prime) \right].$$

**Infer $L$ and $R$:**

$$\ln q(L_v^\prime) = E_{\phi \xi L} \left[p(L, R, \xi, C, Z, \beta, \pi, \alpha, \lambda, \tau, \gamma, X) \right] + \text{const}$$

$$= \sum\nabla_{j,t,k} \chi_{ijt}^v \left[ -\frac{1}{2} E[\xi_k] E \left[ (x_{ijt} - L_i^v R_j)^2 \right] \right]$$

$$- \frac{1}{2} \lambda_{i}^v \cdot \Lambda_{i}^v \cdot L_i^v + \text{const},$$

where $\Lambda_i^v = \text{diag}(E[\xi^v])$. Taking exponential of both sides of Eq. (31), and normalizing the result, we obtain the posterior distribution of $L_v^\prime$:

$$q(L_v^\prime | \mu_v^\prime, \Sigma_v^\prime) = \mathcal{N}(L_v^\prime | \mu_v^\prime, \Sigma_v^\prime),$$

where

$$\Sigma_v^\prime = \left( \sum\nabla_{j,t,k} \chi_{ijt}^v \left[ E[\xi_k] E[R_j R_j^T] + \Lambda_{i}^v \right] \right)^{-1},$$

$$\mu_v^\prime = \sum\nabla_{j,t,k} \chi_{ijt}^v \left[ E[\xi_k] x_{ijt} E[R_j] \right].$$

Similarly, each column of $R$ is also a Gaussian distribution, i.e.,

$$q(R_j | \mu_j, \Sigma_j) = \mathcal{N}(R_j | \mu_j, \Sigma_j),$$

where

$$\Sigma_j = \left( \sum\nabla_{v,i,t,k} \chi_{ijt}^v \left[ E[\xi_k] E[L_i^v L_i^v^T] + \Lambda_{i}^v \right] \right)^{-1},$$

$$\mu_j = \sum\nabla_{v,i,t,k} \chi_{ijt}^v \chi_{ijt}^v \left[ E[\xi_k] x_{ijt} E[L_i^v] \right].$$

and $\Lambda^v = \text{diag}(E[\tau])$.

**Infer $\lambda$ and $\tau$:**

$$\ln q(\lambda_v^\prime) = E_{\phi \xi L} \left[p(L, R, \xi, C, Z, \beta, \pi, \alpha, \lambda, \tau, \gamma, X) \right] + \text{const}$$

$$= \left( \frac{d}{2} + a_0 - 1 \right) \ln \lambda_v^\prime - \frac{1}{2} \left[ E \left[ L_v^\prime R_v^T \right] + b_0 \right] \lambda_v^\prime + \text{const},$$

Thus, we can get the following Updated equations:

$$q(\lambda_v^\prime | a_v^\prime, b_v^\prime) = \text{Gam}(\lambda_v^\prime | a_v^\prime, b_v^\prime),$$

where

$$a_v^\prime = \frac{d}{2} + a_0, \quad b_v^\prime = \frac{1}{2} E \left[ L_v^\prime R_v^T \right] + b_0.$$
TABLE I
RRSE COMPARISON OF NIID-MSL AND TWO BASELINE METHODS ON CMU MULTI-PIE FACE DATASETS WITHOUT ANY SYNTHETIC NOISE. THE BEST RESULTS IN EACH EXPERIMENT ARE HIGHLIGHTED IN RED.

<table>
<thead>
<tr>
<th>Index</th>
<th>Methods</th>
<th>RRSE</th>
<th>RRAE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline 1</td>
<td>Baseline 2</td>
<td>NIID-MSL</td>
</tr>
<tr>
<td>RRSE</td>
<td>0.0121</td>
<td>0.0106</td>
<td>0.0092</td>
</tr>
<tr>
<td>RRAE</td>
<td>0.0680</td>
<td>0.0680</td>
<td>0.0671</td>
</tr>
</tbody>
</table>

TABLE II
F-MEASURE VALUE OF NIID-MSL AND TWO BASELINE METHODS ON WALLFLOWER DATASET. THE BEST RESULTS IN EACH EXPERIMENT ARE HIGHLIGHTED IN RED.

<table>
<thead>
<tr>
<th>Video</th>
<th>Methods</th>
<th>Bootstrapping</th>
<th>Camouflage</th>
<th>Aperture</th>
<th>SwitchLight</th>
<th>TimeOfDay</th>
<th>WavingTrees</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline 1</td>
<td>Baseline 2</td>
<td>NIID-MSL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRSE</td>
<td>0.7326</td>
<td>0.7325</td>
<td>0.7326</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRAE</td>
<td>0.7205</td>
<td>0.7239</td>
<td>0.7413</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRSE</td>
<td>0.9593</td>
<td>0.9592</td>
<td>0.9593</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRAE</td>
<td>0.6825</td>
<td>0.6804</td>
<td>0.6852</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRSE</td>
<td>0.7641</td>
<td>0.7681</td>
<td>0.7594</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRSE</td>
<td>0.7180</td>
<td>0.6647</td>
<td>0.9119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RRSE</td>
<td>0.7628</td>
<td>0.7548</td>
<td>0.7982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DPGMM for all the views (complex but i.i.d.) while the second one different DPGMM for each view of data (complex, non-identical, but independent). Since the latent subspace modeling part of these two baselines are the same with the NIID-MSL as shown in Eq. (3), we only list the noise modeling part of them as follows.

Baseline 1:
\[ \xi_k \sim \text{Gam}(e_0, f_0), \quad e_{ij}^v \sim \mathcal{N}(0, (\xi_{z_{ij}}^v)^{-1}) , \] \hspace{1cm} (58a)
\[ z_{ij}^v \sim \text{Multi}(\pi), \quad \pi_k = \pi_k' \prod_{l} (1 - \pi_k'), \] \hspace{1cm} (58b)
\[ \pi_k' \sim \text{Beta}(1, \gamma), \quad \gamma \sim \text{Gam}(m_0, n_0). \] \hspace{1cm} (58c)

Baseline 2:
\[ \xi_k^v \sim \text{Gam}(e_0, f_0), \quad e_{ij}^v \sim \mathcal{N}(0, (\xi_{z_{ij}}^v)^{-1}) , \] \hspace{1cm} (59a)
\[ z_{ij}^v \sim \text{Multi}(\pi^v), \quad \pi_k^v = \pi_k^v' \prod_{l} (1 - \pi_k^v'), \] \hspace{1cm} (59b)
\[ \pi_k^v' \sim \text{Beta}(1, \gamma^v), \quad \gamma^v \sim \text{Gam}(m_0, n_0). \] \hspace{1cm} (59c)

B. Experimental Results

We compare our proposed NIID-MSL methods with two baselines in Eq. (58) and Eq. (59) to validate the effectiveness of our non-identical and non-independent assumptions on the noise of multi-view data. And some experiments were carried on the real face image recovery (‘No noise’ case of Table III of main text) and real application of foreground detection on RGB data (part C of Section VI), because they can be more representative of the characteristics of noise in practical multi-view data.

Theoretically, Baseline 1 and Baseline 2 are both special cases of our proposed NIID-MSL. As shown in Fig. 2 of the main text, the NIID-MSL degeneratied into Baseline 1 when the MoGs in each view of the second-level all share the same Gaussian components from the first-level. On the contrary, if they do not share any same Gaussian component, the NIID-MSL is equivalent to Baseline 2. The Table I and Table II list the quantitative comparison of RRSE and RRAE in face image recovery and F-Measure in foreground detection experiments, respectively. It is easy to see that NIID-MSL obtains the best or the second best performance in most of the cases, which validates the above theoretical analysis experimentally.

REFERENCES