Graph Wasserstein Correlation Analysis for Movie Retrieval

Xueya Zhang, Tong Zhang, Xiaobin Hong, Zhen Cui, and Jian Yang

Key Lab of Intelligent Perception and Systems for High-Dimensional Information of Ministry of Education, School of Computer Science and Engineering, Nanjing University of Science and Technology
{zhangxueya,tong.zhang,xbhong,zhen.cui,csjyang}@njust.edu.cn

1 Supplementary Material

The detailed derivation processes of $\Sigma_1, (\Sigma_1\Sigma_2)^{1/2}, \mathcal{D}$ and $\sum_{m=1}^{M}\mathcal{D}^{(m)}$ in Eqn (10), Eqn (26) and Eqn (29).

$$
\begin{aligned}
\Sigma_1 &= \frac{1}{n_1}(\widetilde{\mathbf{x}}_1-\mu_1)^\top(\widetilde{\mathbf{x}}_1-\mu_1) && (1)\\
&= \frac{1}{n_1}(\widetilde{\mathbf{x}}_1-\frac{1}{n_1}\mathbf{1}^\top\widetilde{\mathbf{x}}_1)^\top(\widetilde{\mathbf{x}}_1-\frac{1}{n_1}\mathbf{1}^\top\widetilde{\mathbf{x}}_1) && (2)\\
&= \frac{1}{n_1}(\mathbf{L}_1^k\mathbf{X}_1\mathbf{w}_1-\frac{1}{n_1}\mathbf{1}\mathbf{1}^\top\mathbf{L}_1^k\mathbf{X}_1\mathbf{w}_1)^\top(\widetilde{\mathbf{x}}_1-\frac{1}{n_1}\mathbf{1}^\top\widetilde{\mathbf{x}}_1) && (3)\\
&= \frac{1}{n_1}\mathbf{w}_1^\top\mathbf{X}_1^\top(\mathbf{L}_1^k-\frac{1}{n_1}\mathbf{1}\mathbf{1}^\top\mathbf{L}_1^k)^\top(\mathbf{L}_1^k-\frac{1}{n_1}\mathbf{1}\mathbf{1}^\top\mathbf{L}_1^k)\mathbf{X}_1\mathbf{w}_1 && (4)\\
&= \mathbf{w}_1^\top\mathbf{X}_1\mathcal{K}_{\Sigma_1}\mathbf{X}_1\mathbf{w}_1 && (5)
\end{aligned}
$$

$$
\begin{aligned}
(\Sigma_1\Sigma_2)^{1/2} &= (\frac{1}{n_1n_2}(\widetilde{\mathbf{x}}_1')^\top\widetilde{\mathbf{x}}_1'(\widetilde{\mathbf{x}}_2')^\top\widetilde{\mathbf{x}}_2')^{1/2} \geq \frac{1}{\sqrt{n_1n_2}}(\widetilde{\mathbf{x}}_1')^\top\widetilde{\mathbf{x}}_2\\
&= \frac{1}{\sqrt{n_1n_2}}(\widetilde{\mathbf{x}}_1-\mu_1)^\top(\widetilde{\mathbf{x}}_2-\mu_2) && (6)\\
&= \frac{1}{\sqrt{n_1n_2}}\mathbf{w}_1^\top\mathbf{X}_1^\top(\mathbf{L}_1^k-\frac{1}{n_1}\mathbf{1}\mathbf{1}\mathbf{L}_1^k)^\top(\mathbf{L}_2^k-\frac{1}{n_2}\mathbf{1}\mathbf{1}\mathbf{L}_2^k)\mathbf{X}_2\mathbf{w}_2 && (7)\\
&\{\quad \mathcal{K}_{\Sigma_1\Sigma_2} = \frac{1}{\sqrt{n_1n_2}}(\mathbf{L}_1^k-\frac{1}{n_1}\mathbf{1}\mathbf{1}\mathbf{L}_1^k)^\top(\mathbf{L}_2^k-\frac{1}{n_2}\mathbf{1}\mathbf{1}\mathbf{L}_2^k)\quad\} && (8)\\
&= \mathbf{w}_1^\top\mathbf{X}_1^\top\mathcal{K}_{\Sigma_1\Sigma_2}\mathbf{X}_2\mathbf{w}_2 && (9)
\end{aligned}
$$

Xueya Zhang and Tong Zhang have equal contributions.
Corresponding author: zhen.cui@njust.edu.cn.

$$\mathcal{D} = \|\mu_1 - \mu_2\|_2^2 + (\varSigma_1 + \varSigma_2 - 2(\varSigma_1^{1/2}\varSigma_2\varSigma_1^{1/2})^{1/2}) \tag{10}$$

$$\leq \mathbf{w}_1^\top \mathbf{X}_1^\top \mathcal{K}_{\mu_1} \mathbf{X}_1 \mathbf{w}_1 + \mathbf{w}_2^\top \mathbf{X}_2^\top \mathcal{K}_{\mu_2} \mathbf{X}_2 \mathbf{w}_2 \tag{11}$$

$$- 2\mathbf{w}_1^\top \mathbf{X}_1^\top \mathcal{K}_{\mu_1\mu_2} \mathbf{X}_2 \mathbf{w}_2 + \tag{12}$$

$$+ \mathbf{w}_1^\top \mathbf{X}_1^\top \mathcal{K}_{\varSigma_1} \mathbf{X}_1 \mathbf{w}_1 + \mathbf{w}_2^\top \mathbf{X}_2^\top \mathcal{K}_{\varSigma_2} \mathbf{X}_2 \mathbf{w}_2 \tag{13}$$

$$- 2\mathbf{w}_1^\top \mathbf{X}_1^\top \mathcal{K}_{\varSigma_1\varSigma_2} \mathbf{X}_2 \mathbf{w}_2 \tag{14}$$

$$= \mathbf{w}_1^\top \mathbf{X}_1^\top (\mathcal{K}_{\mu_1} + \mathcal{K}_{\varSigma_1}) \mathbf{X}_1 \mathbf{w}_1 \tag{15}$$

$$+ \mathbf{w}_2^\top \mathbf{X}_2^\top (\mathcal{K}_{\mu_2} + \mathcal{K}_{\varSigma_2}) \mathbf{X}_2 \mathbf{w}_2 \tag{16}$$

$$- 2\mathbf{w}_1^\top \mathbf{X}_1^\top (\mathcal{K}_{\mu_1\mu_2} + \mathcal{K}_{\varSigma_2\varSigma_2}) \mathbf{X}_2 \mathbf{w}_2 \tag{17}$$

$$\begin{aligned} \sum_{m=1}^{M} \mathcal{D}^{(m)} = \sum_{m=1}^{M} \{ & \mathbf{w}_1^\top (\mathbf{X}_1^{(m)})^\top (\mathcal{K}_{\mu_1} + \mathcal{K}_{\varSigma_1}) \mathbf{X}_1^{(m)} \mathbf{w}_1 \\ & + \mathbf{w}_2^\top (\mathbf{X}_2^{(m)})^\top (\mathcal{K}_{\mu_2} + \mathcal{K}_{\varSigma_2}) \mathbf{X}_2^{(m)} \mathbf{w}_2 \\ & - 2\mathbf{w}_1^\top (\mathbf{X}_1^{(m)})^\top (\mathcal{K}_{\mu_1\mu_2} + \mathcal{K}_{\varSigma_2\varSigma_2}) \mathbf{X}_2^{(m)} \mathbf{w}_2 \} \end{aligned} \tag{18}$$

$$\{ \quad \mathcal{C}_1 = \sum_{m=1}^{M} (\mathbf{X}_1^{(m)})^\top (\mathcal{K}_{\mu_1} + \mathcal{K}_{\varSigma_1}) \mathbf{X}_1^{(m)} \quad \} \tag{19}$$

$$\{ \quad \mathcal{C}_2 = \sum_{m=1}^{M} (\mathbf{X}_2^{(m)})^\top (\mathcal{K}_{\mu_2} + \mathcal{K}_{\varSigma_2}) \mathbf{X}_2^{(m)} \quad \} \tag{20}$$

$$\{ \quad \mathcal{C}_{12} = \sum_{m=1}^{M} (\mathbf{X}_1^{(m)})^\top (\mathcal{K}_{\mu_1\mu_2} + \mathcal{K}_{\varSigma_1\varSigma_2}) \mathbf{X}_2^{(m)} \quad \} \tag{21}$$

$$= \mathbf{w}_1^\top \mathcal{C}_1 \mathbf{w}_1 + \mathbf{w}_2^\top \mathcal{C}_2 \mathbf{w}_2 - 2\mathbf{w}_1^\top \mathcal{C}_{12} \mathbf{w}_2 \tag{22}$$