Supplementary Material for "Least-Squares Log-Density Gradient Clustering for Riemannian Manifolds"

Mina Ashizawa[1] **Hiroaki Sasaki**[2,3] **Tomoya Sakai**[1] **Masashi Sugiyama**[3,1]
[1]The University of Tokyo, Japan [2]Nara Institute of Science and Technology, Japan [3]RIKEN, Japan

1 Derivation of Eq.(9)

Here, we show the details for the derivation of Eq.(9) from Eq.(8).

For two functions $h, f \in C^\infty(\mathcal{M})$, we have the following relation (Hsu, 2002):

$$\operatorname{div}(h(\boldsymbol{X})\nabla f(\boldsymbol{X})) = \langle \nabla h(\boldsymbol{X}), \nabla f(\boldsymbol{X})\rangle_H + h(\boldsymbol{X})\Delta f(\boldsymbol{X}). \tag{1}$$

Recall $\langle \boldsymbol{X}, \boldsymbol{Y}\rangle_H = \operatorname{tr}(\boldsymbol{X}^\top \boldsymbol{Y})$ and $\boldsymbol{\psi}_l(\boldsymbol{X}) = \nabla\phi_l(\boldsymbol{X}) = -\frac{1}{2\sigma^2}\nabla\delta(\boldsymbol{X}, \boldsymbol{C}_l)^2\phi_l(\boldsymbol{X})$, where $\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^d A_{i,i}$ for a square matrix $\boldsymbol{A} \in \mathbb{R}^{d\times d}$. Using Eq.(1) we have

$$\begin{aligned}
\operatorname{div}(\boldsymbol{\psi}_l(\boldsymbol{X})) &= -\frac{1}{2\sigma^2}\operatorname{div}(\phi_l(\boldsymbol{X})\nabla\delta(\boldsymbol{X}, \boldsymbol{C}_l)^2) \\
&= -\frac{1}{2\sigma^2}\langle \nabla\phi_l(\boldsymbol{X}), \nabla\delta(\boldsymbol{X}, \boldsymbol{C}_l)^2\rangle_H - \frac{1}{2\sigma^2}\phi_l(\boldsymbol{X})\Delta\delta(\boldsymbol{X}, \boldsymbol{C}_l)^2 \\
&= \frac{1}{4\sigma^4}\|\nabla\delta(\boldsymbol{X}, \boldsymbol{C}_l)^2\|^2\phi_l(\boldsymbol{X}) - \frac{1}{2\sigma^2}\phi_\ell(\boldsymbol{X})\sum_{j=1}^d\left[\boldsymbol{P}_{\boldsymbol{X}}\frac{\partial}{\partial X^{(j)}}\nabla\delta(\boldsymbol{X}, \boldsymbol{C}_l)^2\right]^{(j)},
\end{aligned} \tag{2}$$

where $\boldsymbol{P}_{\boldsymbol{X}}$ is the orthogonal projection onto the tangent space $T_{\boldsymbol{X}}\mathcal{M}$ (Hsu, 2002, Theorem 3.1.4). For the Grassmann manifold $\mathcal{G}_{d_1,d_2}$, we have $P_{\boldsymbol{X}} = \boldsymbol{I}_{d_1} - \boldsymbol{X}\boldsymbol{X}^\top$ and $\nabla\delta(\boldsymbol{X}, \boldsymbol{Y})^2 = -2(\boldsymbol{I}_{d_1} - \boldsymbol{X}\boldsymbol{X}^\top)\boldsymbol{Y}\boldsymbol{Y}^\top\boldsymbol{X}$, where $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{G}_{d_1,d_2}$. Plugging these equations into Eq.(2) yields $\widehat{h}_l$ in Eq.(9).

References

E. P. Hsu. *Stochastic Analysis on Manifolds.* American Mathematical Society, 2002.