## 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 \(h, f \in C^\infty(M)\), we have the following relation (Hsu, 2002):

\[
\text{div}(h(X)\nabla f(X)) = \langle \nabla h(X), \nabla f(X) \rangle_H + h(X)\Delta f(X).
\]  

(1)

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

\[
\text{div}(\psi_l(X)) = -\frac{1}{2\sigma^2}\text{div}(\psi_l(X)\nabla \delta(X, C_l)^2)
= -\frac{1}{2\sigma^2}\langle \nabla \phi_l(X), \nabla \delta(X, C_l)^2 \rangle_H - \frac{1}{2\sigma^2}\phi_l(X)\Delta \delta(X, C_l)^2
= \frac{1}{4\sigma^4}\|\nabla \delta(X, C_l)^2\|^2\phi_l(X) - \frac{1}{2\sigma^2}\phi_l(X)\sum_{j=1}^d \left[ P_X \frac{\partial}{\partial X(j)} \nabla \delta(X, C_l)^2 \right]^{(j)},
\]

(2)

where \(P_X\) is the orthogonal projection onto the tangent space \(T_XM\) (Hsu, 2002, Theorem 3.1.4). For the Grassmann manifold \(G_{d_1, d_2}\), we have \(P_X = I_{d_1} - XX^\top\) and \(\nabla \delta(X, Y)^2 = -2(I_{d_1} - XX^\top)YY^\top X\), where \(X, Y \in G_{d_1, d_2}\). Plugging these equations into Eq. (2) yields \(\hat{h}_l\) in Eq. (9).

### References