
Supplementary Material for “Manifold-based Similarity Adaptation for Label Propagation”

Masayuki Karasuyama and Hiroshi Mamitsuka
Bioinformatics Center, Institute for Chemical Research
Kyoto University
Gokasyo, Uji, Kyoto, Japan
{karasuyama, mami}@kuicr.kyoto-u.ac.jp

1 Proof of Theorem 1

Theorem 1. Suppose \mathbf{x}_i can be approximated by its neighbors as follows

$$\mathbf{x}_i = \frac{1}{D_{ii}} \sum_{j \sim i} W_{ij} \mathbf{x}_j + \mathbf{e}_i, \quad (1)$$

where $\mathbf{e}_i \in \mathbb{R}^p$ represents an approximation error. Then, the same adjacency matrix reconstructs the output $y_i \in \mathbb{R}$ with the following error:

$$y_i - \frac{1}{D_{ii}} \sum_{j \sim i} W_{ij} y_j = \mathbf{J} \mathbf{e}_i + O(\|\delta \boldsymbol{\tau}_i\|_2^2) + O(\varepsilon_x + \varepsilon_y), \quad (2)$$

where $\mathbf{J} = \frac{\partial h(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} \left(\frac{\partial g(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} \right)^+$ with superscript $+$ indicates pseudoinverse, and $\delta \boldsymbol{\tau}_i = \max_j (\|\boldsymbol{\tau}_i - \boldsymbol{\tau}_j\|_2^2)$.

Proof. Let $\beta_j = W_{ij}/D_{ii}$ (Note that then $\sum_{j \sim i} \beta_j = 1$). Assuming that g is smooth enough, we obtain the following first-order Taylor expansion at $\boldsymbol{\tau}_i$ for the right hand side of (1).

$$\begin{aligned} \mathbf{x}_i &= \sum_{j \sim i} \beta_j \left(g(\boldsymbol{\tau}_i) + \frac{\partial g(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} (\boldsymbol{\tau}_j - \boldsymbol{\tau}_i) + O(\|\boldsymbol{\tau}_j - \boldsymbol{\tau}_i\|_2^2) \right) \\ &\quad + \mathbf{e}_i + O(\varepsilon_x), \end{aligned}$$

Arranging this equation, we obtain

$$\frac{\partial g(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} \sum_{j \sim i} \beta_j (\boldsymbol{\tau}_j - \boldsymbol{\tau}_i) = -\mathbf{e}_i + O(\|\delta \boldsymbol{\tau}_i\|_2^2) + O(\varepsilon_x).$$

If the Jacobian matrix $\frac{\partial g(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top}$ has full column rank, we obtain

$$\begin{aligned} \sum_{j \sim i} \beta_j (\boldsymbol{\tau}_j - \boldsymbol{\tau}_i) &= - \left(\frac{\partial g(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} \right)^+ \mathbf{e}_i + O(\|\delta \boldsymbol{\tau}_i\|_2^2) \\ &\quad + O(\varepsilon_x). \end{aligned} \quad (3)$$

On the other hand, we can see

$$\begin{aligned}
\sum_{j \sim i} \beta_j y_j &= \sum_{j \sim i} \beta_j \left(h(\boldsymbol{\tau}_i) + \frac{\partial h(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} (\boldsymbol{\tau}_j - \boldsymbol{\tau}_i) \right. \\
&\quad \left. + O(\|\boldsymbol{\tau}_j - \boldsymbol{\tau}_i\|_2^2) \right) + O(\varepsilon_y) \\
&= y_i + \frac{\partial h(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} \sum_{j \sim i} \beta_j (\boldsymbol{\tau}_j - \boldsymbol{\tau}_i) \\
&\quad + O(\|\delta \boldsymbol{\tau}_i\|_2^2) + O(\varepsilon_y)
\end{aligned} \tag{4}$$

Substituting (3) into (4), we obtain

$$\begin{aligned}
y_i - \sum_{j \sim i} \beta_j y_j &= \frac{\partial h(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} \left(\frac{\partial g(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} \right)^\top \mathbf{e}_i \\
&\quad + O(\|\delta \boldsymbol{\tau}_i\|_2^2) + O(\varepsilon_x + \varepsilon_y).
\end{aligned}$$

□

2 Theorem 1 for Normalized Laplacian

Theorem 2. Suppose that \mathbf{x}_i can be approximated by its neighbors as follows

$$\mathbf{x}_i = \sum_{j \sim i} \frac{W_{ij}}{\sqrt{D_{ii} D_{jj}}} \mathbf{x}_j + \mathbf{e}_i, \tag{5}$$

where $\mathbf{e}_i \in \mathbb{R}^p$ represents an approximation error. Then, the same adjacency matrix reconstructs the output $y_i \in \mathbb{R}$ with the following error:

$$\begin{aligned}
y_i - \sum_{j \sim i} \frac{W_{ij}}{\sqrt{D_{ii} D_{jj}}} y_j &= (1 - \sum_{j \sim i} \gamma_j) (h(\boldsymbol{\tau}_i) + \mathbf{J} g(\boldsymbol{\tau}_i)) \\
&\quad + \mathbf{J} \mathbf{e}_i + O(\|\delta \boldsymbol{\tau}_i\|_2^2) + O(\varepsilon_x + \varepsilon_y),
\end{aligned} \tag{6}$$

where

$$\gamma_j = \frac{W_{ij}}{\sqrt{D_{ii} D_{jj}}}.$$

Proof. The proof is almost the same as Theorem 1. However, the sum of the coefficients γ_j (corresponding to β_j in Theorem 1) cannot be 1. Applying the same Taylor expansion to the right hand side of (5), we obtain

$$\begin{aligned}
\frac{\partial g(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} \sum_{j \sim i} \gamma_j (\boldsymbol{\tau}_j - \boldsymbol{\tau}_i) &= -\mathbf{e}_i + (1 - \sum_{j \sim i} \gamma_j) g(\boldsymbol{\tau}_i) \\
&\quad + O(\|\delta \boldsymbol{\tau}_i\|_2^2) + O(\varepsilon_x).
\end{aligned}$$

On the other hand, applying Taylor expansion to $y_i - \sum_{j \sim i} \gamma_j y_j$, we obtain

$$\begin{aligned}
y_i - \sum_{j \sim i} \gamma_j y_j &= (1 - \sum_{j \sim i} \gamma_j) h(\boldsymbol{\tau}_i) - \frac{\partial h(\boldsymbol{\tau}_i)}{\partial \boldsymbol{\tau}^\top} \sum_{j \sim i} \gamma_j (\boldsymbol{\tau}_j - \boldsymbol{\tau}_i) \\
&\quad + O(\|\delta \boldsymbol{\tau}_i\|_2^2) + O(\varepsilon_y)
\end{aligned}$$

Using the above two equations, we obtain (6). □