Supplementary Material for Generalized Domain-Adaptive Dictionaries

Sumit Shekhar Vishal M. Patel Hien V. Nguyen
Rama Chellappa
University of Maryland, College Park, USA
{sshekha,pvishalm,hien,rama}@umiacs.umd.edu

1. Optimization

The final optimization problem is given as:

$$\{\mathbf{D}^*, \tilde{\mathbf{P}}^*, \tilde{\mathbf{X}}^*\} = \arg\min_{\mathbf{D},\tilde{\mathbf{P}},\tilde{\mathbf{X}}} \mathcal{C}_1(\mathbf{D}, \tilde{\mathbf{P}}, \tilde{\mathbf{X}}) + \lambda \mathcal{C}_2(\tilde{\mathbf{P}})$$
$$\text{s.t. } \mathbf{P_i P_i^T} = \mathbf{I},\ i = 1, \cdots, M \text{ and } \|\tilde{\mathbf{x}}_j\|_1 \leq T_0, \forall j \quad (1)$$

where,

$$\mathcal{C}_1(\mathbf{D}, \tilde{\mathbf{P}}, \tilde{\mathbf{X}}) = \|\tilde{\mathbf{P}}\tilde{\mathbf{Y}} - \mathbf{D}\tilde{\mathbf{X}}\|_F^2 + \mu\|\tilde{\mathbf{P}}\tilde{\mathbf{Y}} - \mathbf{D}\tilde{\mathbf{X}}_{\mathbf{in}}\|_F^2 + \nu\|\mathbf{D}\tilde{\mathbf{X}}_{\mathbf{out}}\|_F^2, \quad (2)$$
$$\mathcal{C}_2(\tilde{\mathbf{P}}) = -\text{trace}((\tilde{\mathbf{P}}\tilde{\mathbf{Y}})(\tilde{\mathbf{P}}\tilde{\mathbf{Y}})^T). \quad (3)$$

1.1. Proposition 1:

There exists an optimal solution $\mathbf{P}_\mathbf{1}^*, \cdots, \mathbf{P}_\mathbf{M}^*, \mathbf{D}^*$ to equation (1), which has the following form:

$$\mathbf{P}_\mathbf{i}^* = (\mathbf{Y_i A_i})^\mathbf{T}\ \forall\, i = 1, \cdots, M \quad (4)$$
$$\mathbf{D}^* = \tilde{\mathbf{P}}^* \tilde{\mathbf{Y}} \tilde{\mathbf{B}} \quad (5)$$

Proof:

Form for $\mathbf{D}^*$: First we will show the form for $\mathbf{D}^*$. We can decompose $\mathbf{D}^*$ into two orthogonal components as follows:

$$\mathbf{D}^* = \mathbf{D}_\parallel + \mathbf{D}_\perp \quad (6)$$
$$\text{where, } \mathbf{D}_\parallel = (\tilde{\mathbf{P}}\tilde{\mathbf{Y}})\tilde{\mathbf{B}},\ \mathbf{D}_\perp^\mathbf{T}(\tilde{\mathbf{P}}\tilde{\mathbf{Y}}) = \mathbf{0} \quad (7)$$

for some $\mathbf{B} \in \mathbb{R}^{\sum_{i=1}^M N_i \times K}$. Substituting the value of $\mathbf{D}^*$ into the value of $\mathcal{C}_1(\mathbf{D}, \tilde{\mathbf{P}}, \tilde{\mathbf{X}})$, we get for the three terms of $\mathcal{C}_1$, ignoring the multiplicative constants μ, ν:

$$\textbf{First Term} = \text{trace}((\tilde{\mathbf{P}}\tilde{\mathbf{Y}} - \mathbf{D}\tilde{\mathbf{X}})^T(\tilde{\mathbf{P}}\tilde{\mathbf{Y}} - \mathbf{D}\tilde{\mathbf{X}}))$$
$$= \text{trace}(\tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{P}}^\mathbf{T}\tilde{\mathbf{P}}\tilde{\mathbf{Y}} + \tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{P}}^\mathbf{T}\mathbf{D}_\parallel\tilde{\mathbf{X}} + \tilde{\mathbf{X}}^\mathbf{T}\mathbf{D}_\parallel^\mathbf{T}\mathbf{D}_\parallel\tilde{\mathbf{X}} + \tilde{\mathbf{X}}^\mathbf{T}\mathbf{D}_\perp^\mathbf{T}\mathbf{D}_\perp\tilde{\mathbf{X}})$$
$$\geq \text{trace}(\tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{P}}^\mathbf{T}\tilde{\mathbf{P}}\tilde{\mathbf{Y}} + \tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{P}}^\mathbf{T}\mathbf{D}_\parallel\tilde{\mathbf{X}} + \tilde{\mathbf{X}}^\mathbf{T}\mathbf{D}_\parallel^\mathbf{T}\mathbf{D}_\parallel\tilde{\mathbf{X}}). \quad (8)$$

$$\textbf{Second Term} = \text{trace}((\tilde{\mathbf{P}}\tilde{\mathbf{Y}} - \mathbf{D}\tilde{\mathbf{X}}_\mathbf{in})^T(\tilde{\mathbf{P}}\tilde{\mathbf{Y}} - \mathbf{D}\tilde{\mathbf{X}}_\mathbf{in}))$$
$$= \text{trace}(\tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{P}}^\mathbf{T}\tilde{\mathbf{P}}\tilde{\mathbf{Y}} + \tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{P}}^\mathbf{T}\mathbf{D}_\parallel\tilde{\mathbf{X}}_\mathbf{in} + \tilde{\mathbf{X}}_\mathbf{in}^\mathbf{T}\mathbf{D}_\parallel^\mathbf{T}\mathbf{D}_\parallel\tilde{\mathbf{X}}_\mathbf{in} + \tilde{\mathbf{X}}_\mathbf{in}^\mathbf{T}\mathbf{D}_\perp^\mathbf{T}\mathbf{D}_\perp\tilde{\mathbf{X}}_\mathbf{in})$$
$$\geq \text{trace}(\tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{P}}^\mathbf{T}\tilde{\mathbf{P}}\tilde{\mathbf{Y}} + \tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{P}}^\mathbf{T}\mathbf{D}_\parallel\tilde{\mathbf{X}}_\mathbf{in} + \tilde{\mathbf{X}}_\mathbf{in}^\mathbf{T}\mathbf{D}_\parallel^\mathbf{T}\mathbf{D}_\parallel\tilde{\mathbf{X}}_\mathbf{in}). \quad (9)$$

$$\textbf{Third Term} = \text{trace}(\mathbf{D}\tilde{\mathbf{X}}_\mathbf{out})^\mathbf{T}(\mathbf{D}\tilde{\mathbf{X}}_\mathbf{out}))$$
$$= \text{trace}(\tilde{\mathbf{X}}_\mathbf{out}^\mathbf{T}\mathbf{D}_\parallel^\mathbf{T}\mathbf{D}_\parallel\tilde{\mathbf{X}}_\mathbf{out} + \tilde{\mathbf{X}}_\mathbf{out}^\mathbf{T}\mathbf{D}_\perp^\mathbf{T}\mathbf{D}_\perp\tilde{\mathbf{X}}_\mathbf{out}).$$
$$\geq \text{trace}(\tilde{\mathbf{X}}_\mathbf{out}^\mathbf{T}\mathbf{D}_\parallel^\mathbf{T}\mathbf{D}_\parallel\tilde{\mathbf{X}}_\mathbf{out}) \quad (10)$$

The equality is reached when $\mathbf{D}_\perp = \mathbf{0}$. Hence, the form of $\mathbf{D}^*$ is:

$$\mathbf{D}^* = \tilde{\mathbf{P}}\tilde{\mathbf{Y}}\tilde{\mathbf{B}}.$$

Form for $\mathbf{P}_\mathbf{i}^*$: For each $i = 1, \cdots, M$, $\mathbf{P}_\mathbf{i}^*$ can be decomposed as:

$$\mathbf{P}_\mathbf{i}^* = \mathbf{P}_{\parallel,\mathbf{i}} + \mathbf{P}_{\perp,\mathbf{i}} \quad (11)$$
$$\text{where, } \mathbf{P}_{\parallel,\mathbf{i}} = (\mathbf{Y_i A_i})^\mathbf{T}, \mathbf{P}_{\perp,\mathbf{i}}\mathbf{Y_i} = \mathbf{0}. \quad (12)$$

Let $\tilde{\mathbf{P}}_\parallel = [\mathbf{P}_{\parallel,\mathbf{1}}, \cdots, \mathbf{P}_{\parallel,\mathbf{M}}]$ and $\tilde{\mathbf{P}}_\perp = [\mathbf{P}_{\perp,\mathbf{1}}, \cdots, \mathbf{P}_{\perp,\mathbf{M}}]$. Substituting the value for $\mathbf{D}^*$ into cost terms, we can write the terms of $\mathcal{C}_1$ as:

$$\textbf{First Term} = \|\tilde{\mathbf{P}}^*\tilde{\mathbf{Y}}(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})\|_F^2$$
$$= \|(\tilde{\mathbf{P}}_\parallel + \tilde{\mathbf{P}}_\perp)\tilde{\mathbf{Y}}(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})\|_F^2$$
$$= \|\tilde{\mathbf{P}}_\parallel\tilde{\mathbf{Y}}(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})\|_F^2$$
$$= \text{trace}(\tilde{\mathbf{P}}_\parallel\tilde{\mathbf{Y}}(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})^\mathbf{T}\tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{P}}_\parallel^\mathbf{T}). \quad (13)$$

$$\textbf{Second Term} = \|\tilde{\mathbf{P}}^*\tilde{\mathbf{Y}}(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_\mathbf{in})\|_F^2$$
$$= \|(\tilde{\mathbf{P}}_\parallel + \tilde{\mathbf{P}}_\perp)\tilde{\mathbf{Y}}(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_\mathbf{in})\|_F^2$$
$$= \|\tilde{\mathbf{P}}_\parallel\tilde{\mathbf{Y}}(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_\mathbf{in})\|_F^2$$
$$= \text{trace}(\tilde{\mathbf{P}}_\parallel\tilde{\mathbf{Y}}(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_\mathbf{in})(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_\mathbf{in})^\mathbf{T}\tilde{\mathbf{Y}}^\mathbf{T}\tilde{\mathbf{P}}_\parallel^\mathbf{T}). \quad (14)$$

$$\begin{aligned}
&\textbf{Third Term} = \|\tilde{\mathbf{P}}^* \tilde{\mathbf{Y}}(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})\|_F^2 \\
&= \|(\tilde{\mathbf{P}}_{\|} + \tilde{\mathbf{P}}_{\perp})\tilde{\mathbf{Y}}(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})\|_F^2 \\
&= \|\tilde{\mathbf{P}}_{\|}\tilde{\mathbf{Y}}(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})\|_F^2 \\
&= \text{trace}(\tilde{\mathbf{P}}_{\|}\tilde{\mathbf{Y}}(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})^{\mathbf{T}}\tilde{\mathbf{Y}}^{\mathbf{T}}\tilde{\mathbf{P}}_{\|}^{\mathbf{T}}). \qquad (15)
\end{aligned}$$

The cost term, $\mathcal{C}_2$ can be written as:

$$\begin{aligned}
&\mathcal{C}_2(\tilde{\mathbf{P}}) = -\text{trace}((\tilde{\mathbf{P}}\tilde{\mathbf{Y}})(\tilde{\mathbf{P}}\tilde{\mathbf{Y}})^T) \\
&= -\text{trace}(((\tilde{\mathbf{P}}_{\|} + \tilde{\mathbf{P}}_{\perp})\tilde{\mathbf{Y}})((\tilde{\mathbf{P}}_{\|} + \tilde{\mathbf{P}}_{\perp})\tilde{\mathbf{Y}})^T) \\
&= -\text{trace}((\tilde{\mathbf{P}}_{\|}\tilde{\mathbf{Y}})(\tilde{\mathbf{P}}_{\|}\tilde{\mathbf{Y}})^T). \qquad (16)
\end{aligned}$$

Putting all the terms together, the overall objective function becomes:

$$\begin{aligned}
&\text{trace}(\tilde{\mathbf{P}}_{\|}\tilde{\mathbf{Y}}((\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})^{\mathbf{T}} + \mu(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{in}}) \\
&(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{in}})^{\mathbf{T}} + \nu(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})^{\mathbf{T}} - \lambda\mathbf{I})\tilde{\mathbf{Y}}^{\mathbf{T}}\tilde{\mathbf{P}}_{\|}^{\mathbf{T}}) \\
&= \text{trace}(\tilde{\mathbf{A}}_{\mathbf{T}}\tilde{\mathbf{K}}((\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})^{\mathbf{T}} + \mu(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{in}}) \\
&(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{in}})^{\mathbf{T}} + \nu(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})^{\mathbf{T}} - \lambda\mathbf{I})\tilde{\mathbf{K}}\tilde{\mathbf{A}}). \qquad (17)
\end{aligned}$$

It can be seen that from equation (17), that the cost function is independent of $\mathbf{P}_{\perp,\mathbf{i}}$, hence it can be safely set to be $\mathbf{0}$. Hence,

$$\mathbf{P}_{\mathbf{i}}^* = (\mathbf{Y}_{\mathbf{i}}\mathbf{A}_{\mathbf{i}})^{\mathbf{T}}.$$

2. Updating $\tilde{\mathrm{A}}$

Using Proposition 1, optimization problem equation (1) becomes:

$$\begin{aligned}
&\{\tilde{\mathbf{A}}^*, \tilde{\mathbf{B}}^*, \mathbf{X}^*\} = \underset{\tilde{\mathbf{A}},\tilde{\mathbf{B}},\tilde{\mathbf{X}}}{\arg\min}\ \mathcal{C}_1(\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{X}}) + \lambda\mathcal{C}_2(\tilde{\mathbf{A}}) \\
&\text{s.t. } \mathbf{A}_{\mathbf{i}}^{\mathbf{T}}\mathbf{K}_{\mathbf{i}}\mathbf{A}_{\mathbf{i}} = \mathbf{I},\ i = 1, \cdots, M \text{ and } \|\tilde{\mathbf{x}}_j\|_1 \leq T_0, \forall j. \qquad (18)
\end{aligned}$$

Here, we assume that $(\tilde{\mathbf{B}}, \tilde{\mathbf{X}})$ are fixed. Then, the optimization for $\tilde{\mathbf{A}}$ can be solved efficiently. We have the following proposition.

3. Proposition 2:

The optimal solution of equation (18) when $(\tilde{\mathbf{B}}, \tilde{\mathbf{X}})$ *are fixed is:*

$$\begin{aligned}
&\{\mathbf{G}^*\} = \underset{\mathbf{G}}{\arg\min}\ \text{trace}[\mathbf{G}^{\mathbf{T}}\mathbf{H}\mathbf{G}] \\
&\textit{s.t. } \mathbf{G}_{\mathbf{i}}^{\mathbf{T}}\mathbf{G}_{\mathbf{i}} = \mathbf{I}\ \forall\ i\ = 1, \cdots, M \qquad (19)
\end{aligned}$$

where,

$$\begin{aligned}
&\mathbf{H} = \mathbf{S}^{\frac{1}{2}}\mathbf{V}^{\mathbf{T}}((\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})^{\mathbf{T}} + \mu(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{in}}) \\
&(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{in}})^{\mathbf{T}} + \nu(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})^{\mathbf{T}} - \lambda\mathbf{I})\mathbf{V}\mathbf{S}^{\frac{1}{2}} \qquad (20)
\end{aligned}$$

Proof:

Let,

$$\tilde{\mathbf{K}} = \mathbf{V}\mathbf{S}\mathbf{V}^{\mathbf{T}},$$

$$\begin{aligned}
&\tilde{\mathbf{H}} = \mathbf{S}^{\frac{1}{2}}\mathbf{V}^{\mathbf{T}}((\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}})^{\mathbf{T}} + \\
&\mu(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{in}})(\mathbf{I} - \tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{in}})^{\mathbf{T}} + \nu(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})(\tilde{\mathbf{B}}\tilde{\mathbf{X}}_{\mathbf{out}})^{\mathbf{T}} \\
&-\lambda\mathbf{I})\mathbf{V}\mathbf{S}^{\frac{1}{2}},
\end{aligned}$$

and

$$\mathbf{G} = \mathbf{S}^{\frac{1}{2}}\mathbf{V}^{\mathbf{T}}\tilde{\mathbf{A}}.$$

Substituting into equation (17), we get the required form of the optimization.