Supplementary Material for: Coupled End-to-end Transfer Learning with Generalized Fisher Information

Shixing Chen[1] Caojin Zhang[2] Ming Dong[1]
[1]Department of Computer Science [2]Department of Mathematics
Wayne State University
{schen, czhang, mdong}@wayne.edu

A. Theoretical Analysis

A.1. Proof of lemma 1

Lemma 1 *If* $E_P|S-c_{src}| \le C$, $C>0$, $E_Q|c_{src}-c_{tgt}| \le \lambda_1$, $\sup_f |E_P f - E_{T_1\circ S\in P} f| \le \lambda_2$, *for any* $f \in P,Q$, $\lambda_1, \lambda_2 > 0$, *then there exists some constant* $C>0$, *for any measurable function* $f \in P,Q$, $f>0$,

$$\begin{aligned} E_Q|T_2-c_{tgt}| \quad &\le C + |E_{T_1\circ S\in P} f - E_Q f| + E_Q|T_2-S| \\ &+2\sup_f |E_{T_1\circ S(x)\in Q} f - E_{x\in Q} f| \\ &+2\sup_f |E_{T_1\circ T_2(x)\in Q} f - E_{x\in Q} f| \end{aligned} \tag{1}$$

Proof

$$\begin{aligned} E_Q|T_2-c_{tgt}| \quad &\le E_P|S-c_{src}| + |E_Q|T_2-c_{tgt}| - E_Q|S-c_{src}|| \\ &+|E_Q|S-c_{src}| - E_P|S-c_{src}|| \end{aligned} \tag{2}$$

Using triangle inequality,

$$\begin{aligned} |E_Q|T_2-c_{tgt}| - E_Q|S-c_{src}|| \quad &\le |E_Q|T_2-c_{tgt}-S+c_{src}|| \\ &\le E_Q|T_2-S| + E_Q|c_{tgt}-c_{src}| \end{aligned} \tag{3}$$

It follows:

$$\begin{aligned} E_Q|T_2-c_{tgt}| \le \quad & E_P|S-c_{src}| + E_Q|T_2-S| + E_Q|c_{tgt}-c_{src}| \\ & |E_Q|S-c_{src}| - E_P|S-c_{src}|| \\ & \le C + \lambda_1 + E_Q|T_2-S| \\ & +|E_Q|S-c_{src}| - E_P|S-c_{src}|| \end{aligned} \tag{4}$$

Then, we focus on the analysis of the last term,

$$\begin{aligned} |E_Q|S-c_{src}| \quad - \quad & E_P|S-c_{src}|| \\ & \le \sup_f |E_Q f - E_P f| \\ & \le \lambda_2 + \sup_f |E_{T_1\circ S\in P} f - E_{T_1\circ S\in Q} f| \\ & + \sup_f |E_{T_1\circ S\in Q} f - E_{T_1\circ T_2\in Q} f| \\ & + \sup_f |E_{T_1\circ T_2\in Q} f - E_Q f| \end{aligned} \tag{5}$$

Since

$$\begin{aligned} & |E_{T_1\circ S\in Q} f - E_{T_1\circ T_2\in Q} f| \\ & \le |E_{T_1\circ S\in Q} f - E_Q f| + |E_{T_1\circ T_2\in Q} f - E_Q f| \end{aligned} \tag{6}$$

and

$$\begin{aligned} & |E_{T_1\circ S\in P} f - E_{T_1\circ S\in Q} f| \\ & \le |E_{T_1\circ S\in P} f - E_Q f| + |E_{T_1\circ S\in Q} f - E_Q f| \end{aligned} \tag{7}$$

Combining (5), (6), (7), we get

$$\begin{aligned} |E_Q|S-c_{src}| \quad - \quad & E_P|S-c_{src}|| \\ & \le \lambda_2 + \sup_f |E_{T_1\circ S\in P} f - E_Q f| \\ & +2\sup_f |E_{T_1\circ S\in Q} f - E_Q f| \\ & +2\sup_f |E_{T_1\circ T_2\in Q} f - E_Q f| \end{aligned} \tag{8}$$

Substituting (8) into (4), the desired result follows.

A.2. Proof of lemma 2

Lemma 2 *Assume* $E_{T_1\circ S\in P}|c_{src}-c_{tgt}| \le C$, $E_{T_1\circ S\in Q}|c_{src}-c_{tgt}| \le C$, $E_{T_1\circ S\in P}|c_{tgt}| \le C$, $E_{T_1\circ S\in Q}|c_{src}| < C$, *for some* $C>0$, *then there exists some constant* $C>0$, *such that for any measurable function* $f>0$, *and* $f \in P,Q$,

$$\begin{aligned} \sup_f |E_{T_1\circ S\in P} f - E_Q f| \quad &\le C + E_{T_1\circ S\in P}|S-c_{src}| \\ &+\sup_f |E_{T_1\circ S\in Q} f - E_Q f| \\ &+E_Q|T_2-S| \end{aligned} \tag{9}$$

Proof

Using the definition of $\sup_f$, for S, there exists some constant $C>0$, such that:

$$\sup_f |E_{T_1\circ S\in P} f - E_Q f| \le |E_{T_1\circ S\in P} S - E_Q S| + C \tag{10}$$

It follows:

$$\begin{aligned} \sup_f |E_{T_1\circ S\in P} f - E_Q f| \le |E_{T_1\circ S\in P} S - E_{T_1\circ S\in P}|c_{src}-c_{tgt}| \\ +E_{T_1\circ S\in Q}|c_{src}-c_{tgt}| - E_Q S| + C \end{aligned} \tag{11}$$

Using triangle inequality, we get

$$\begin{aligned} \sup_f |E_{T_1\circ S\in P} f - E_Q f| \quad &\le E_{T_1\circ S\in P}|S-c_{src}| + C \\ &+|E_{T_1\circ S\in Q}|c_{src}-c_{tgt}| - E_Q S| + C \\ &\le C + E_{T_1\circ S\in P}|S-c_{src}| \\ &+|E_{T_1\circ S\in Q} c_{tgt} - E_Q T_2| + E_Q|T_2-S| \\ &\le C + E_{T_1\circ S\in P}|S-c_{src}| \\ &+\sup_f |E_{T_1\circ S\in Q} f - E_Q f| + E_Q|T_2-S| \end{aligned} \tag{12}$$