Supplementary Material of "Selecting Effective Triplet Contrastive Loss for Domain Alignment"

Changshi Li[1] and Zhijian Yang[2]

[1] School of Computer Science, Wuhan University, Wuhan, 430072, P.R. China
[2] School of Mathematics and Statistics and Hubei Key Laboratory of Computational Science, Wuhan University, Wuhan, 430072, P.R. China
zjyang.math@whu.edu.cn

Theorem 1. *Let $\boldsymbol{z}_i$ and $\boldsymbol{z}_j$ be the representations of two positive samples. Then, the cosine similarity of any two positive pairs satisfies*

$$\hat{\boldsymbol{z}}_i^T \hat{\boldsymbol{z}}_j \geq \frac{4}{\tau} \log \frac{|S_-|}{(e^{\mathcal{L}_{CL}} - 1)} - 7, \tag{1}$$

Theorem 2. *Let c be the number of classes, $\{w_1, w_2, \cdots, w_c\}$ be a set of anchors, where w_p is the anchor of representations from the p-th class. Then the contrastive loss $\mathcal{L}_{SCL}$ satisfies*

$$\log(1 + \frac{|S_-|}{e^{2\tau}}) \leq \mathcal{L}_{SCL} \leq 2\tau \sqrt{4 - \frac{2}{\tau} \log \frac{|S_-|}{e^{\mathcal{L}_W} - 1}} + \mathcal{L}_W, \tag{2}$$

where the triplet contrastive loss $\mathcal{L}_W$ can be any one of $\mathcal{L}_w$, $\mathcal{L}_s$ and $\mathcal{L}_{ws}$. They are defined as follows respectively

$$\mathcal{L}_w = -\log \frac{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau}}{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{j \neq p} W_j e^{\hat{\boldsymbol{w}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau}}, \tag{3}$$

$$\mathcal{L}_s = -\log \frac{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau}}{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{q \in Q} e^{\hat{\boldsymbol{z}}_q^T \hat{\boldsymbol{z}}_i \cdot \tau}}, \tag{4}$$

$$\mathcal{L}_{ws} = -\log \frac{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau}}{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{j \neq p} e^{\hat{\boldsymbol{w}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{q \in Q} e^{\hat{\boldsymbol{z}}_q^T \hat{\boldsymbol{z}}_i \cdot \tau}}, \tag{5}$$

where W_j is the number of representations that share the same anchor w_j.

Proof. Let $\boldsymbol{z}_i$ be the representation, $\boldsymbol{w}_p$ be the positive anchor of $\boldsymbol{z}_i$, S_- be a set of negative pairs. We denote the triplet contrastive loss as

$$\mathcal{L}_{CL} = -\log(\frac{e^{s_+ \cdot \tau}}{e^{s_+ \cdot \tau} + \sum_{s_- \in S_-} e^{s_- \cdot \tau}}), \tag{6}$$

where $s_+ = \hat{\boldsymbol{z}}_i^T \hat{\boldsymbol{w}}_p$ and $s_- = \hat{\boldsymbol{z}}_i^T \hat{\boldsymbol{w}}_q$. Then s_+ satisfies

$$s_+ = \frac{1}{\tau} \log(\frac{\sum_{s_- \in S_-} e^{s_- \cdot \tau}}{e^{\mathcal{L}_{CL}} - 1}) \geq \frac{1}{\tau} \log \frac{|S_-|}{(e^{\mathcal{L}_{CL}} - 1)} - 1. \tag{7}$$

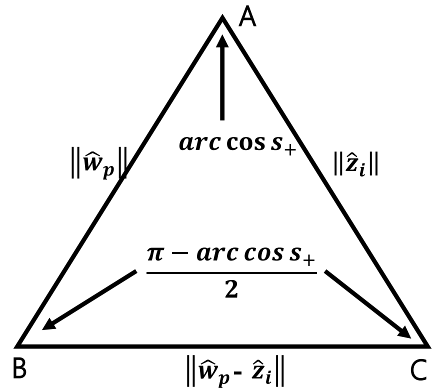

Fig. 1. Isosceles triangle is composed by $\boldsymbol{a}_p$ and $\boldsymbol{z}_i$. A, B and C are the angle of isosceles triangle.

Then we consider $\|\hat{\boldsymbol{w}}_p - \hat{\boldsymbol{z}}_i\|_2$ as follows. Because $\hat{\boldsymbol{w}}_p$ and $\hat{\boldsymbol{z}}_i$ are normalized vectors, $\hat{\boldsymbol{w}}_p$ and $\hat{\boldsymbol{z}}_i$ can form the edge of an isosceles triangle, as shown in fig. 1 . According to vector operations, the base of isosceles triangle is $\hat{\boldsymbol{w}}_p - \hat{\boldsymbol{z}}_i$. Angle A is $\arccos \hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i = \arccos s_+$. Angle B and C are $\frac{\pi - \arccos s_+}{2}$ according to the sum of the internal angles in an isosceles triangle. So the length of the base $\|\hat{\boldsymbol{w}}_p - \hat{\boldsymbol{z}}_i\|_2$ satisfies

$$\begin{aligned} &\|\hat{\boldsymbol{w}}_p - \hat{\boldsymbol{z}}_i\|_2 = \\ &\|\hat{\boldsymbol{w}}_p\|_2 \cos B + \|\hat{\boldsymbol{z}}_i\|_2 \cos C \\ &= 2 \cos \frac{\pi - \arccos s_+}{2}. \end{aligned} \tag{8}$$

According to the lower bound of Equation (7), we have

$$\begin{aligned} \|\hat{\boldsymbol{w}}_p - \hat{\boldsymbol{z}}_i\|_2 &\leq 2 \cos(\frac{\pi - \arccos r}{2}) \\ &= 2 \sin(\frac{\arccos r}{2}) \\ &= \sqrt{2 - 2r}, \end{aligned} \tag{9}$$

where $r := \frac{1}{\tau} \log \frac{|S_-|}{(e^{\mathcal{L}_{CL}} - 1)} - 1$. According to the triangle inequality, the distance between any $\hat{\boldsymbol{z}}_i$ and $\hat{\boldsymbol{z}}_j$ from positive pairs satisfies

$$\|\hat{\boldsymbol{z}}_i - \hat{\boldsymbol{z}}_j\|_2 \leq \|\hat{\boldsymbol{w}}_p - \hat{\boldsymbol{z}}_i\|_2 + \|\hat{\boldsymbol{w}}_p - \hat{\boldsymbol{z}}_j\|_2 \leq 2\sqrt{2 - 2r}. \tag{10}$$

Finally, the dot between $\hat{\boldsymbol{z}}_i$ and $\hat{\boldsymbol{z}}_j$ satisfies

$$\hat{\boldsymbol{z}}_i^T \hat{\boldsymbol{z}}_j \geq 4r - 3 = \frac{4}{\tau} \log \frac{|S_-|}{(e^{\mathcal{L}_{CL}} - 1)} - 7, \tag{11}$$

which completes the proof of Theorem 1.

In the next, we will show the relation between contrastive loss and triplet contrastive loss. Let c be the number of classes, $\{w_1, w_2, \cdots, w_c\}$ be a set of anchors, where w_p is the positive anchor of representations from p-th class.

Given contrastive loss

$$\mathcal{L}_{SCL} = -\log\left(\frac{e^{s_+ \cdot \tau}}{e^{s_+ \cdot \tau} + \sum_{s_- \in S_-} e^{s_- \cdot \tau}}\right). \tag{12}$$

Given the triplet contrastive loss

$$\mathcal{L}_W = -\log\left(\frac{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau}}{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{s_- \in S_-} e^{s_- \cdot \tau}}\right). \tag{13}$$

According to Equation (7), we have $\|\hat{\boldsymbol{z}}_j - \hat{\boldsymbol{w}}_p\|_2 \le \sqrt{2-2r}$ and $r = \frac{1}{\tau} \log \frac{|S_-|}{(e^{\mathcal{L}_W} - 1)} - 1$. Since the fact of Cauchy-Schwarz inequality $(\hat{\boldsymbol{z}}_j - \hat{\boldsymbol{w}}_p)^T \hat{\boldsymbol{z}}_i$ or $(\hat{\boldsymbol{w}}_p - \hat{\boldsymbol{z}}_j)^T \hat{\boldsymbol{z}}_i \le \|\hat{\boldsymbol{z}}_i\|_2 \|\hat{\boldsymbol{z}}_j - \hat{\boldsymbol{w}}_p\|_2 \le \sqrt{2-2r} = \delta$, we consider the first upper bound of contrastive loss

$$\begin{aligned}
\mathcal{L}_{SCL} &= -\log \frac{e^{\hat{\boldsymbol{z}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau}}{e^{\hat{\boldsymbol{z}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{q \in Q} e^{\hat{\boldsymbol{z}}_q^T \hat{\boldsymbol{z}}_i \cdot \tau}} \\
&= -\hat{\boldsymbol{z}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau + \log(e^{\hat{\boldsymbol{z}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{q \in Q} e^{\hat{\boldsymbol{z}}_q^T \hat{\boldsymbol{z}}_i \cdot \tau}) \\
&\le -\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau + \log(e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau + \delta\tau} + \sum_{j \ne p} e^{\hat{\boldsymbol{w}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau + \delta\tau}) + \delta\tau \\
&= 2\delta\tau - \log \frac{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau}}{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{j \ne p} W_j e^{\hat{\boldsymbol{w}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau}} \\
&= 2\delta\tau + \mathcal{L}_w,
\end{aligned} \tag{14}$$

where W_j is the number of representations which have the same positive anchor $\boldsymbol{w}_j$. We consider the second upper bound of contrastive loss

$$\begin{aligned}
\mathcal{L}_{SCL} &= -\hat{\boldsymbol{z}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau + \log(e^{\hat{\boldsymbol{z}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{q \in Q} e^{\hat{\boldsymbol{z}}_q^T \hat{\boldsymbol{z}}_i \cdot \tau}) \\
&\le -\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau + \log(e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau + \delta\tau} + \sum_{q \in Q} e^{\hat{\boldsymbol{z}}_q^T \hat{\boldsymbol{z}}_i \cdot \tau + \delta\tau}) + \delta\tau \\
&= 2\delta\tau - \log \frac{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau}}{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{q \in Q} e^{\hat{\boldsymbol{z}}_q^T \hat{\boldsymbol{z}}_i \cdot \tau}} \\
&= 2\delta\tau + \mathcal{L}_s.
\end{aligned} \tag{15}$$

Finally, we have the third triplet contrastive loss

$$\begin{aligned}\mathcal{L}_{SCL} &\leq 2\delta\tau - \log \frac{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau}}{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{q \in Q} e^{\hat{\boldsymbol{z}}_q^T \hat{\boldsymbol{z}}_i \cdot \tau}} \\ &\leq 2\delta\tau - \log \frac{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau}}{e^{\hat{\boldsymbol{w}}_p^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{j \neq p} e^{\hat{\boldsymbol{w}}_j^T \hat{\boldsymbol{z}}_i \cdot \tau} + \sum_{q \in Q} e^{\hat{\boldsymbol{z}}_q^T \hat{\boldsymbol{z}}_i \cdot \tau}} \\ &= 2\delta\tau + \mathcal{L}_{ws}.\end{aligned} \tag{16}$$

Obviously, all of the above equations have the same form, namely $\mathcal{L}_{SCL} \leq 2\delta\tau + \mathcal{L}_W$. The proof for Theorem 2 is completed.