Lemma 0.1. *Given an embedding X and triplet constraints $T = \{(i,j,k) : d(x_i,x_j)^2 < d(x_i,x_k)^2\}$, if $d(x_i,x_j)^2 - d(x_i,x_k)^2 > 1$ whenever X satisfies constraint $(i,j,k) \in T$, then X is not a local minimum of*

$$L(X) = \sum_{(i,j,k)\in T} \max\{0, 1 - (d(x_i,x_j)^2 - d(x_i,x_k)^2)\}.$$

Proof. The main idea is that under the above assumptions, the unviolated triplet constraints contribute nothing to the hinge loss since the gaps between square distances are large. We can find a direction to move in by scaling the original embedding so that the hinge loss of the violated triplet constraints goes down while mainting the large gaps of squared differences in the constraints that are satisfied.

Given a proposed embedding X such that $d(x_i,x_j)^2 - d(x_i,x_k)^2 > 1$ for all triplet constraints $(i,j,k) \in T$ that X satisfies, let

$$S_X = \{(i,j,k) \in T : d(x_i,x_j)^2 - d(x_i,x_k)^2 > 1\}$$

correspond to the set of satisfied triplet constraints under X and let

$$U_X = \{(i,j,k) \in T : d(x_i,x_j)^2 - d(x_i,x_k)^2 < 0\}$$

correspond to the set of unsatisfied triplet constraints under X.
Let $\beta = \min_{(i,j,k)\in S_X}\{d(x_i,x_j)^2 - d(x_i,x_k)^2\} > 1$. Now consider $Y_\alpha = \sqrt{\alpha}X$ for $\frac{1}{\beta} < \alpha < 1$.

Then $d(y_i,y_j)^2 = \alpha d(x_i,x_j)^2$. Therefore, for $(i,j,k) \in S_X$,

$$d(y_i,y_j)^2 - d(y_i,y_k)^2 = \alpha(d(x_i,x_j)^2 - d(x_i,x_k)^2) > \frac{d(x_i,x_j)^2 - d(x_i,x_k)^2}{\beta} > 1$$

by definition of α and β, so

$$\max\{0, 1 - (d(y_i,y_j)^2 - d(y_i,y_k)^2)\} = 0. \tag{1}$$

Similarly, if $(i,j,k) \in U_X$ then

$$d(y_i,y_j)^2 - d(y_i,y_k)^2 = \alpha^2(d(x_i,x_j)^2 - d(x_i,x_k)^2) < 0$$

and since $\alpha < 1$ and $d(x_i,x_j)^2 - d(x_i,x_k)^2 < 0$,

$$d(y_i,y_j)^2 - d(y_i,y_k)^2 = \alpha(d(x_i,x_j)^2 - d(x_i,x_k)^2) > d(x_i,x_j)^2 - d(x_i,x_k)^2,$$

so

$$\begin{aligned}\max\{0, 1 - (d(y_i,y_j)^2 - d(y_i,y_k)^2)\} &= 1 - (d(y_i,y_j)^2 - d(y_i,y_k)^2) \\ &< 1 - (d(y_i,y_j)^2 - d(y_i,y_k)^2) \\ &= \max\{0, 1 - (d(x_i,x_j)^2 - d(x_i,x_k)^2)\},\end{aligned}$$

Hence,

$$\max\{0, 1 - (d(y_i,y_j)^2 - d(y_i,y_k)^2)\} < \max\{0, 1 - (d(x_i,x_j)^2 - d(x_i,x_k)^2)\}. \tag{2}$$

Furthermore, we see that $U_X = U_{Y_\alpha}$ and $S_X = S_{Y_\alpha}$ and $U_{Y_\alpha} \sqcup S_{Y_\alpha} = T$

Recall, we want to show that X is not a local minimum. This means we need to show that there is some sufficiently close embedding Y where $L(Y) < L(X)$.

To formalize this, let $\epsilon > 0$ and consider $B_{X,\epsilon} = \{Y \in \mathbb{R}^{nd} : ||X - Y||_2 < \epsilon\}$ where we view X is a $n \times d$ vector.
Now let $1 > \alpha > \max\{1 - \frac{\epsilon}{||X||_2}, \frac{1}{\beta}\}$. First, it's not hard to see with algebra that such an α exists and $Y_\alpha \in B_{X,\epsilon}$.

Finally, we will show that $L(Y_\alpha) < L(X)$ establishing that X is not a local minimum.

$$\begin{aligned}
L(Y_\alpha) &= \sum_{(i,j,k)\in T} \max\{0, 1 - (d(y_i, y_j)^2 - d(y_i, y_k)^2)\} \\
&= \sum_{(i,j,k)\in U_{Y_\alpha}} \max\{0, 1 - (d(y_i, y_j)^2 - d(y_i, y_k)^2)\} + \sum_{(i,j,k)\in S_{Y_\alpha}} \max\{0, 1 - (d(y_i, y_j)^2 - d(y_i, y_k)^2)\} \\
&= \sum_{(i,j,k)\in U_{Y_\alpha}} \max\{0, 1 - (d(y_i, y_j)^2 - d(y_i, y_k)^2)\} \text{ by Equation 1} \\
&< \sum_{(i,j,k)\in U_X} \max\{0, 1 - (d(x_i, x_j)^2 - d(x_i, x_k)^2)\} \text{ by Equation 2} \\
&= \sum_{(i,j,k)\in T} \max\{0, 1 - (d(x_i, x_j)^2 - d(x_i, x_k)^2)\} \\
&= \sum_{(i,j,k)\in U_X} \max\{0, 1 - (d(x_i, x_j)^2 - d(x_i, x_k)^2)\} + \sum_{(i,j,k)\in S_X} \max\{0, 1 - (d(x_i, x_j)^2 - d(x_i, x_k)^2)\} \\
&\text{by assumption on the squared distance gaps} \\
&= L(X)
\end{aligned}$$

$\square$