# CSCI 145 Problem 17 Solution: Can All Negative Pairs Be Orthogonal?

## Solution

### 1.

The $(i, j)$ entry of $\mathbf{Z}\mathbf{Z}^\top$ is $\mathbf{z}_i^\top \mathbf{z}_j$. Since each embedding has unit norm, its diagonal entries are $\mathbf{z}_i^\top \mathbf{z}_i = 1$. Expanding the Frobenius norm gives

$$\begin{aligned}\mathcal{L} &= \sum_{i=1}^{n} (\|\mathbf{z}_i\|_2^2 - 1)^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} (\mathbf{z}_i^\top \mathbf{z}_j)^2 \\ &= \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} (\mathbf{z}_i^\top \mathbf{z}_j)^2.\end{aligned}$$

In the drawing, $\mathbf{z}_1$ and $\mathbf{z}_2$ are perpendicular, while each makes an obtuse angle with $\mathbf{z}_3$. Only negative-pair errors remain. Each unordered pair appears twice, once as $(i, j)$ and once as $(j, i)$, because the loss sums over all matrix entries.

### 2.

Using symmetry and the identity $\|\mathbf{A}\|_F^2 = \operatorname{tr}(\mathbf{A}^\top \mathbf{A})$, we obtain

$$\begin{aligned}\mathcal{L} &= \operatorname{tr}\big((\mathbf{Z}\mathbf{Z}^\top - \mathbf{I}_n)^2\big) \\ &= \operatorname{tr}(\mathbf{Z}\mathbf{Z}^\top \mathbf{Z}\mathbf{Z}^\top) - 2\operatorname{tr}(\mathbf{Z}\mathbf{Z}^\top) + n.\end{aligned}$$

The diagonal entries of $\mathbf{Z}\mathbf{Z}^\top$ are one, so its trace is $n$. The cyclic trace identity moves the first factor $\mathbf{Z}$ to the end:

$$\begin{aligned}\operatorname{tr}(\mathbf{Z}\mathbf{Z}^\top \mathbf{Z}\mathbf{Z}^\top) &= \operatorname{tr}(\mathbf{Z}^\top \mathbf{Z}\mathbf{Z}^\top \mathbf{Z}) \\ &= \|\mathbf{Z}^\top \mathbf{Z}\|_F^2.\end{aligned}$$

Here $\mathbf{Z}^\top \mathbf{Z}$ is symmetric and $k \times k$. Substituting these identities gives

$$\mathcal{L} = \|\mathbf{Z}^\top \mathbf{Z}\|_F^2 - 2n + n = \|\mathbf{Z}^\top \mathbf{Z}\|_F^2 - n.$$

## 3.

The matrix $\mathbf{Z}^\top \mathbf{Z}$ is symmetric and positive semidefinite. Its orthogonal eigendecomposition shows that its trace is the sum of its $k$ eigenvalues and its squared Frobenius norm is the sum of their squares. Zero eigenvalues are included if the matrix is singular. Thus

$$\sum_{j=1}^{k} \lambda_j = \operatorname{tr}(\mathbf{Z}^\top \mathbf{Z}) = \operatorname{tr}(\mathbf{Z}\mathbf{Z}^\top) = n.$$

By part 2 and the squared-Frobenius identity,

$$\mathcal{L} = \sum_{j=1}^{k} \lambda_j^2 - n.$$

Apply Cauchy–Schwarz to the vectors $(\lambda_1, \ldots, \lambda_k)$ and $(1, \ldots, 1)$:

$$n^2 = \left( \sum_{j=1}^{k} \lambda_j \cdot 1 \right)^2 \leq \left( \sum_{j=1}^{k} \lambda_j^2 \right) \left( \sum_{j=1}^{k} 1^2 \right) = k \sum_{j=1}^{k} \lambda_j^2.$$

Consequently,

$$\mathcal{L} \geq \frac{n^2}{k} - n = \frac{n(n-k)}{k}.$$

This is positive because $n > k$. Even a nonlinear encoder cannot place more than $k$ unit vectors in $\mathbb{R}^k$ with all distinct pairs perpendicular. The bound quantifies the unavoidable total error for this squared-similarity loss, even though the positive pairs agree exactly.