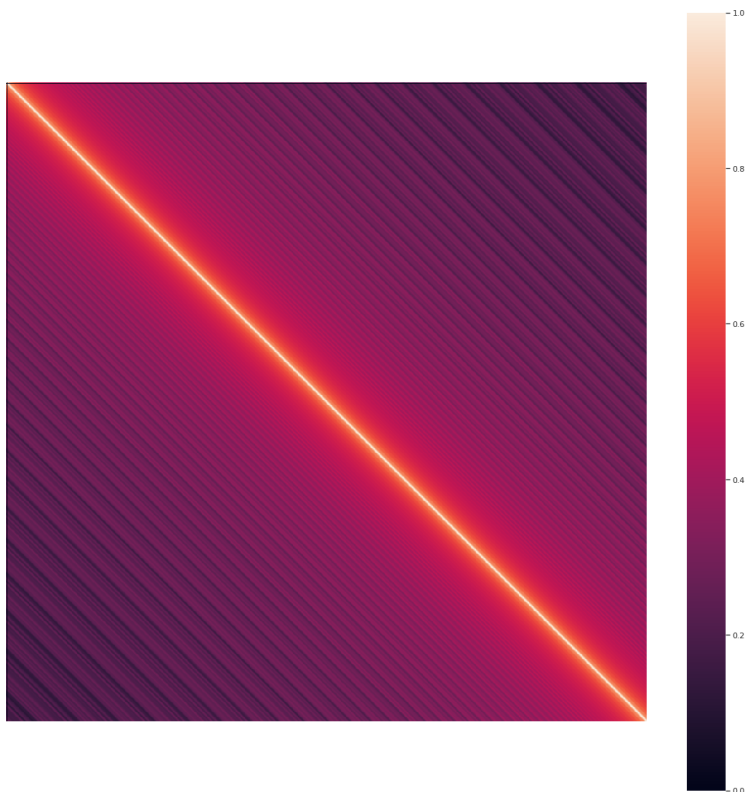


1. Problem 1: Visualize Positional Embedding

Ans: Positional embedding represents location information of a word inside a sentence. Nearby embeddings are more similar than those far away. The figure below shows the fact that diagonal elements always have higher value than others, which means nearby embeddings are always similar to each other due to their locations.



2. Visualize Gradient Norm

