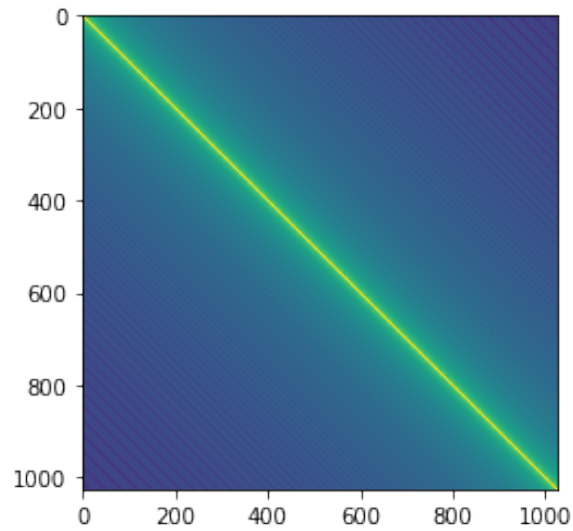


Problem 1: Visualize Positional Embedding

說明：Embedding 中 Positions 越相近的相似度越大，反之，Positions 相距越遠的相似度越小。且斜對角為同位置的相似度，均為 1。



Problem 2: Gradient Explosion

圈出兩處發生梯度爆炸的位置：

