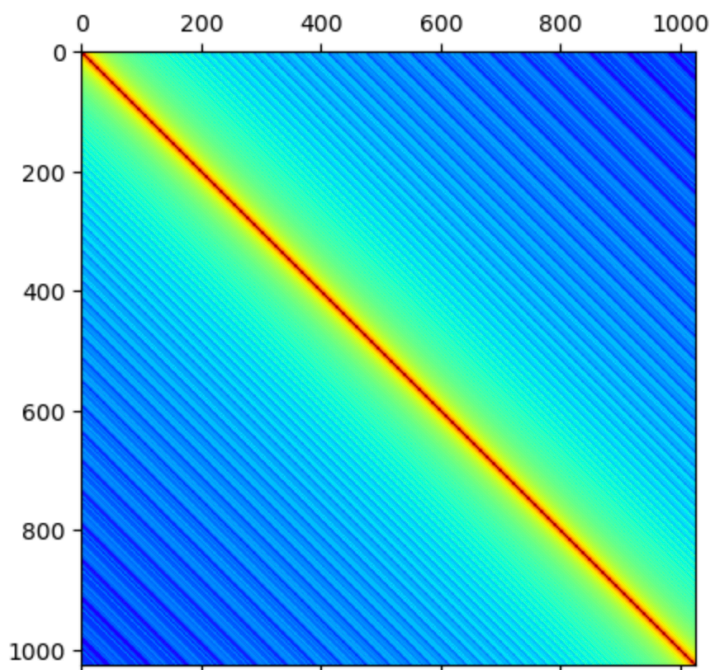# ML HW5 Gradescope
B09602017 EE3 白宗民

Problem 1.

It can be observed that the strongest similarity is on the diagonal, and the similarity is symmetric with respect to the diagonal. Additionally, the similarity at positions equidistant from the diagonal is the same (exhibiting translational invariance).

Size of positional embedding torch.Size([1026, 256])
<Figure size 800x800 with 0 Axes>

Code:

```python
from torch.nn.functional import cosine_similarity as sim
import matplotlib.pyplot as plt
%matplotlib inline
```

```python
pos_emb = model.decoder.embed_positions.weights.cpu().detach()
print('Size of positional embedding', pos_emb.size())
ret = sim(pos_emb.unsqueeze(1), pos_emb, dim=2)
plt.figure(figsize=(8,8))
plt.matshow(ret, cmap='jet')
plt.show()
```

Problem 2.
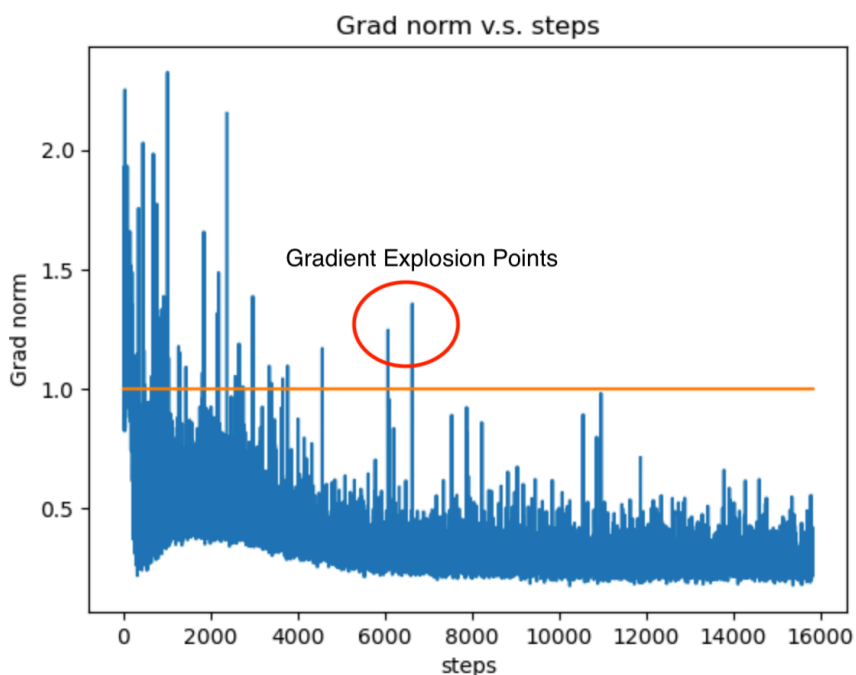
Reference: https://blog.csdn.net/weixin_42369818/article/details/124102920