和自己的positional embedding的相似度最高，距離自己越遠相似度有遞減趨勢。

```python
# reference https://ithelp.ithome.com.tw/articles/10211370

import pandas as pd
import matplotlib.pyplot as plt
pos_emb = model.decoder.embed_positions.weights.cpu().detach()
matrix = [[None for i in range(1026)] for j in range(1026)]
for i in range(1026):
    for j in range(1026):
            matrix[i][j] = F.cosine_similarity(pos_emb[i], pos_emb[j],dim=0).item()
```

```python
plt.imshow(matrix, interpolation='nearest')
plt.show()
plt.plot(mylist)
plt.ylabel("Grad norm")
plt.xlabel("Step")
plt.show()
```