

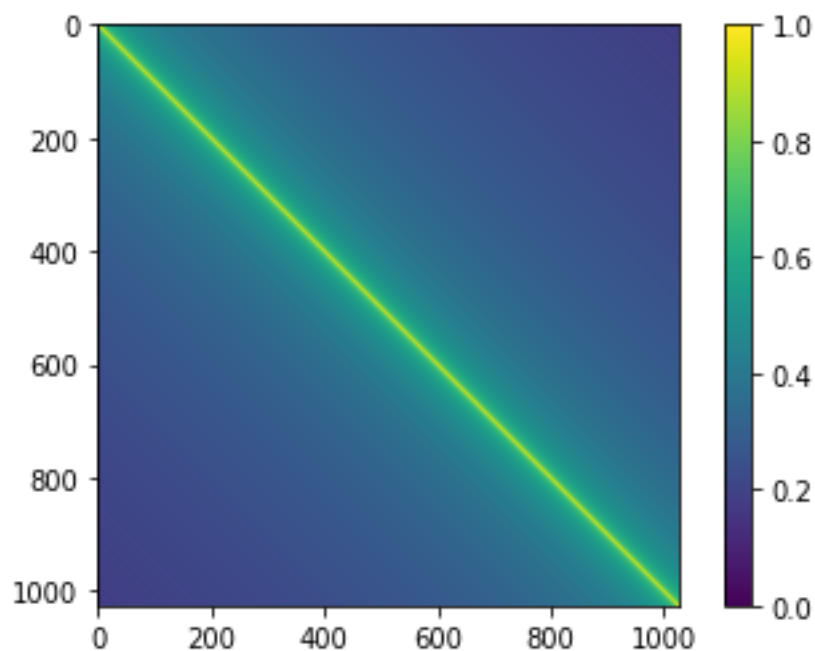
Hw5 Report

R09921026 李育倫

Q1: 由下結果可得知位置彼此相距較近的相關性較高，這也符合預期中所想的。

```
Import torch
Import torch.nn.functional as F
Import matplotlib.pyplot as plt
From tqdm import tqdm
pos_emb = model.decoder.embed_positions.weights.cpu().detach()
print(pos_emb.shape)
pos_emb_list = []
for i in tqdm(range(1026)):
    for j in range(1026):
        output = F.cosine_similarity(pos_emb[i], pos_emb[j], dim=0)
        pos_emb_list.append(output.data.numpy())
pos_emb_numpy = np.array(pos_emb_list).reshape(1026,1026)
print(pos_emb_numpy.shape)
plt.imshow(pos_emb_numpy)
plt.colorbar()
plt.show()
```

```
torch.Size([1026, 1024])
100%|████████████████████████████████████████| 1026/1026 [00:46<00:00, 22.07it/s]
(1026, 1026)
```



Q2: 圈出的地方為超過 `clip_norm = 1.0` 的地方，此時 `clip_grad_norm` function 將會產生作用。

