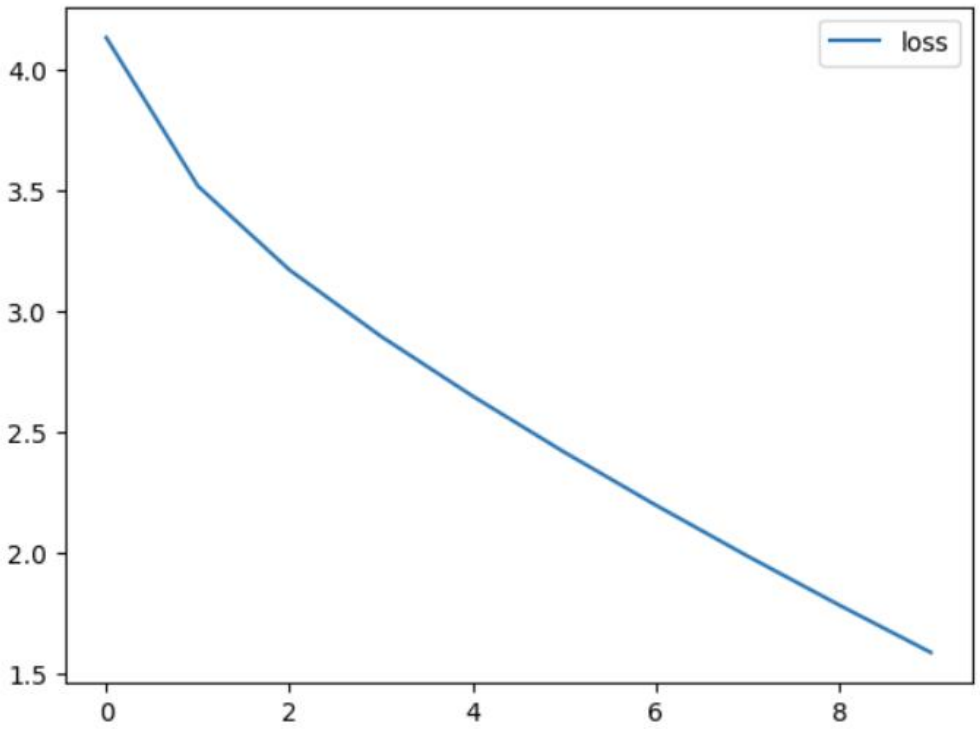
HW 4: Important Figures and Questions

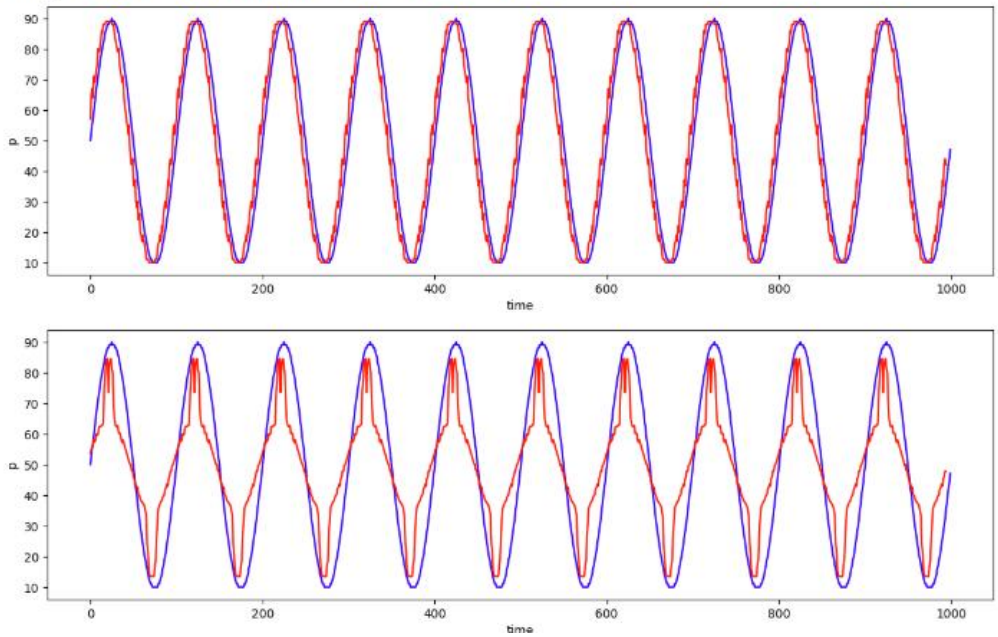
Question 1:

This is the loss

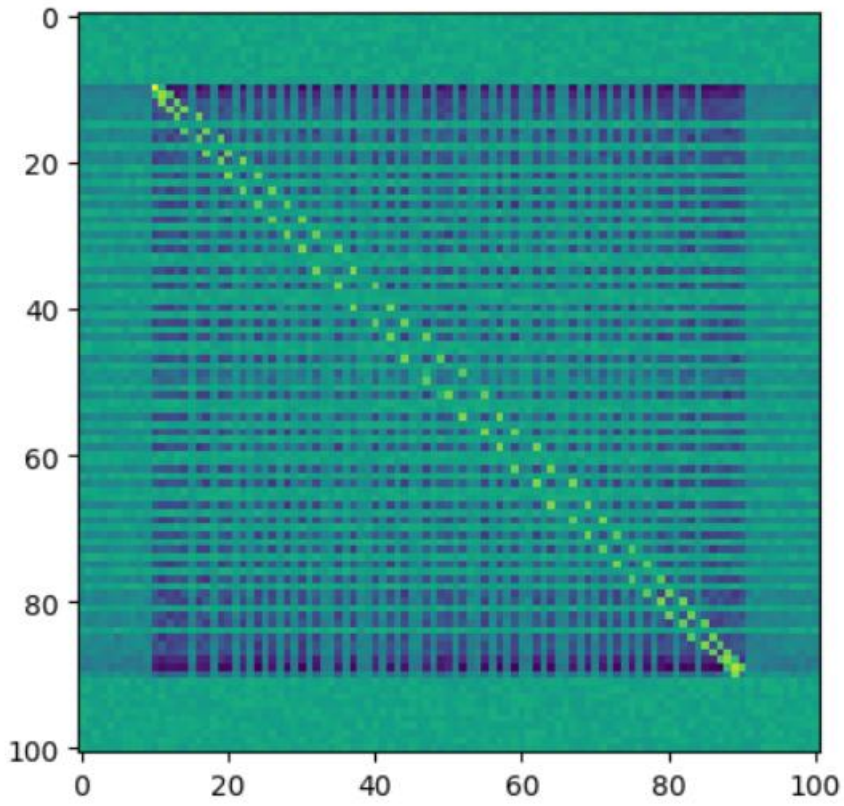


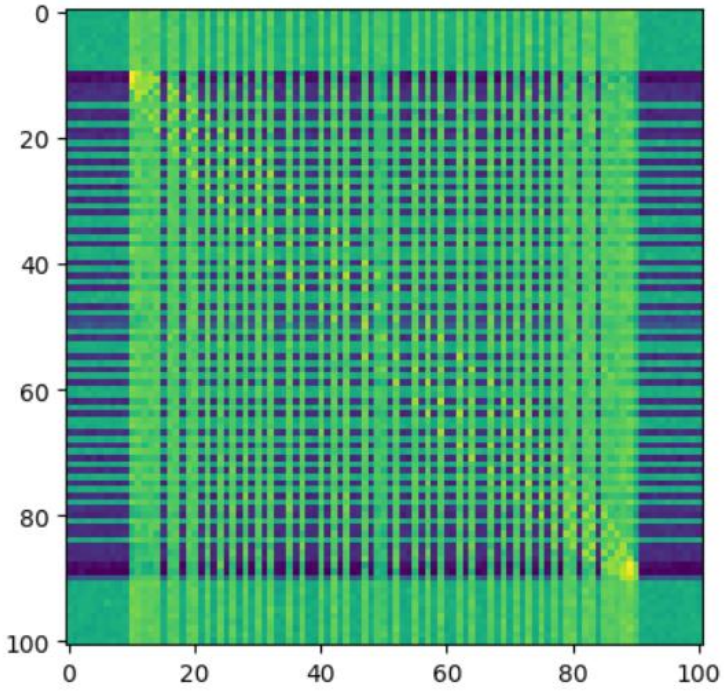
The error and associated graphs are:

```
rms_error ML estimate 240.77998255669013
rms_error MMSE estimate 454.2433218937356
```



The two weights shown as images are:

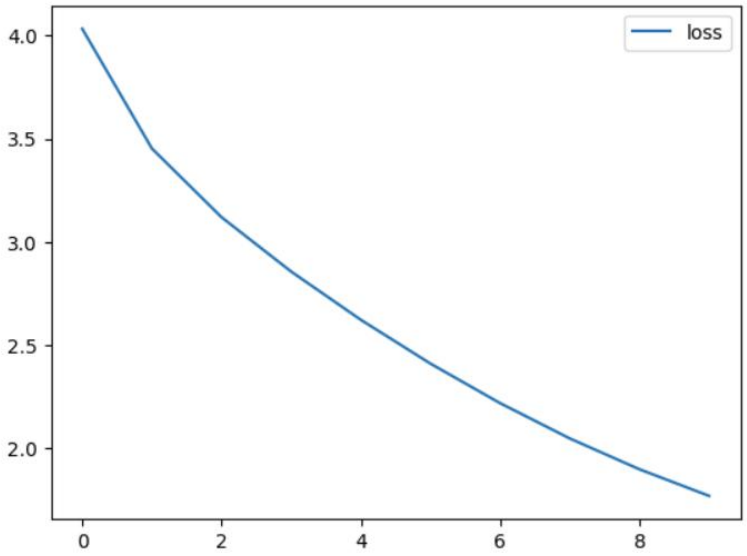




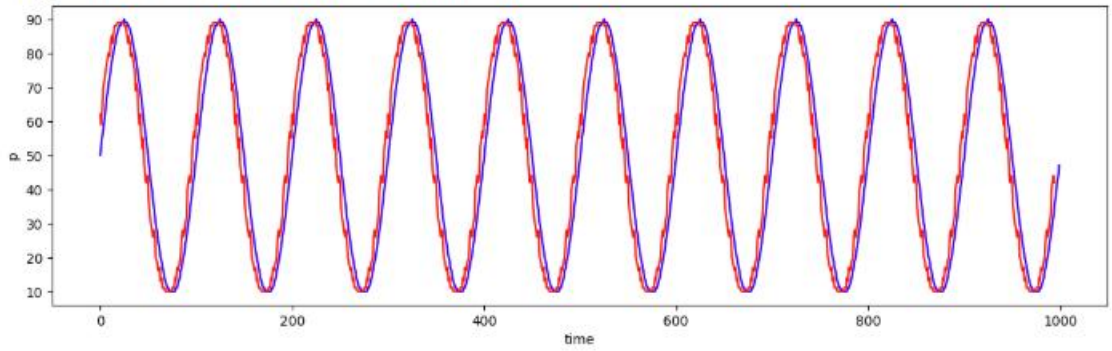
From the above time series given, This is a more predictable timeseries than the one above. What could you do make use of predictability to get better estimates?

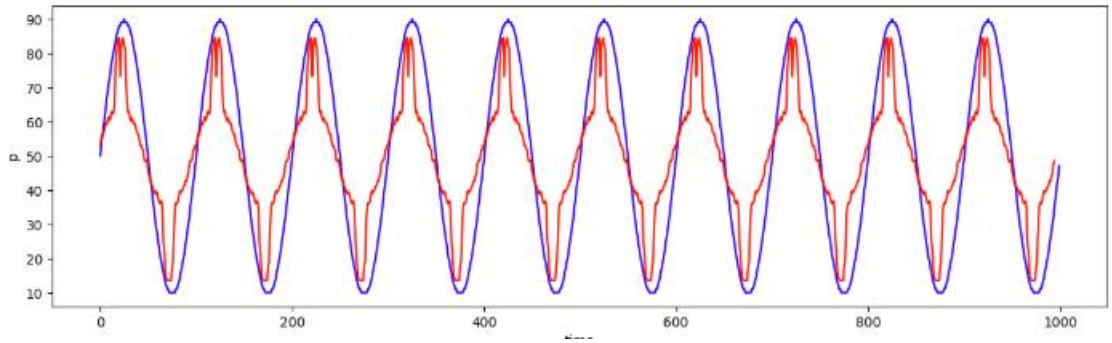
Answer: Because this is a sinusoidal function, we can use a Fourier series model by definition as it is just a sum of sinusoids. Furthermore, since we know the function ahead of time it is deterministic, not random and thus we can tell the network exactly the function to model and use a type of numerical error calculation like MSE to make the model as accurate as possible.

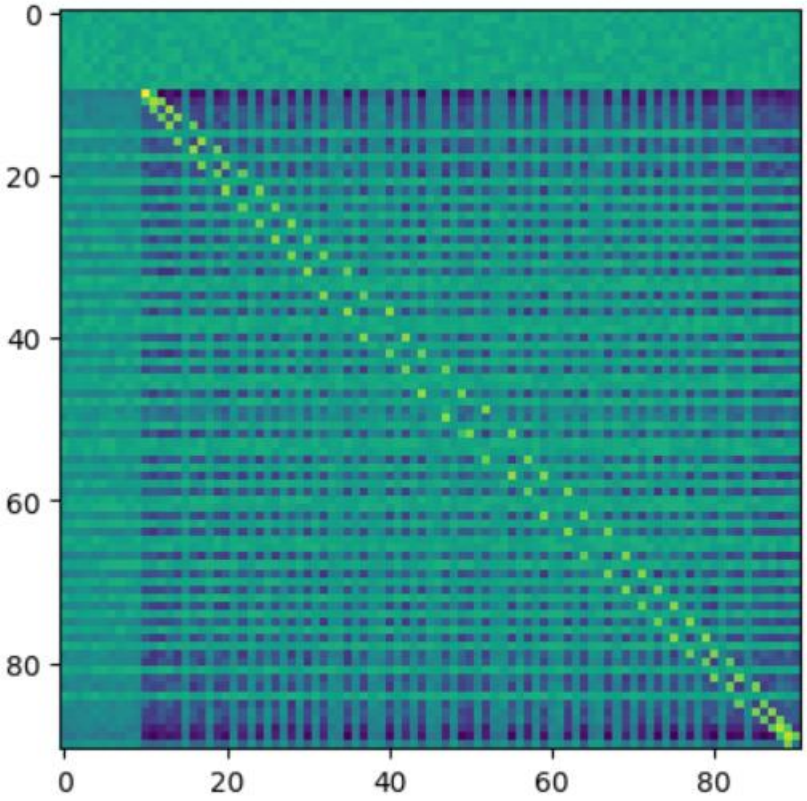
Question 2:

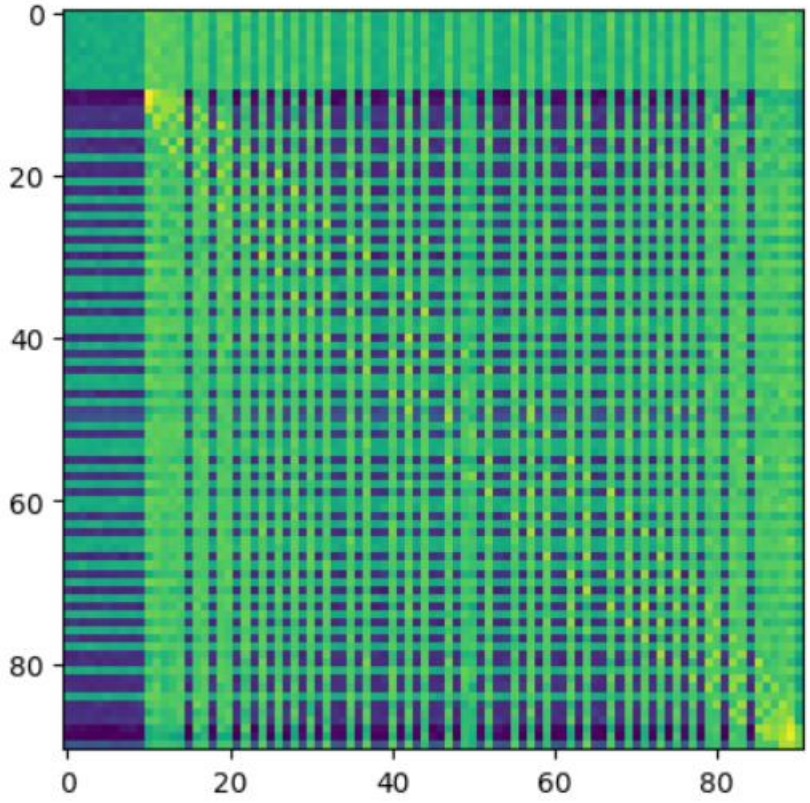


```
rms_error ML estimate 239.47860029656096
rms_error MMSE estimate 454.68023890205416
```





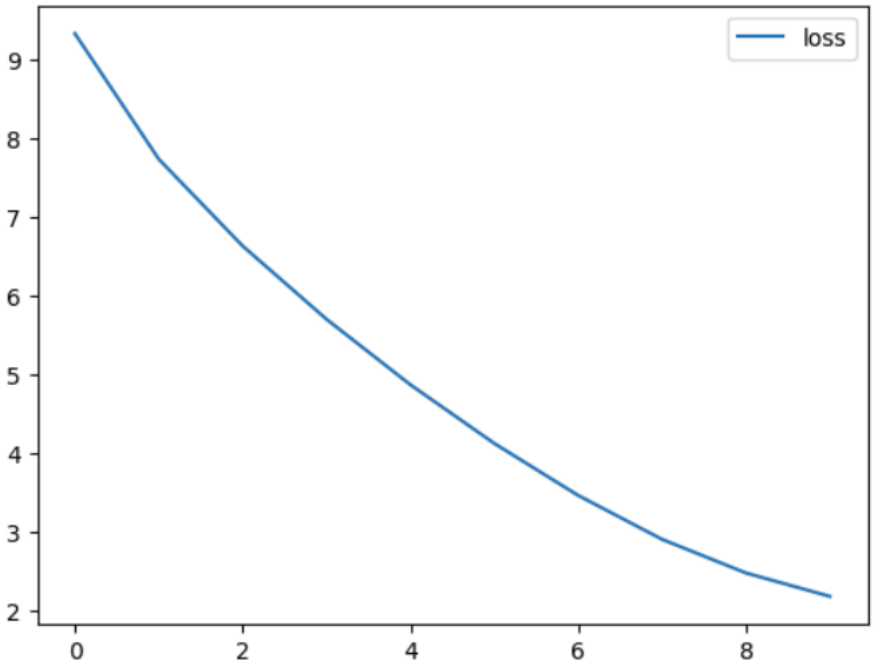




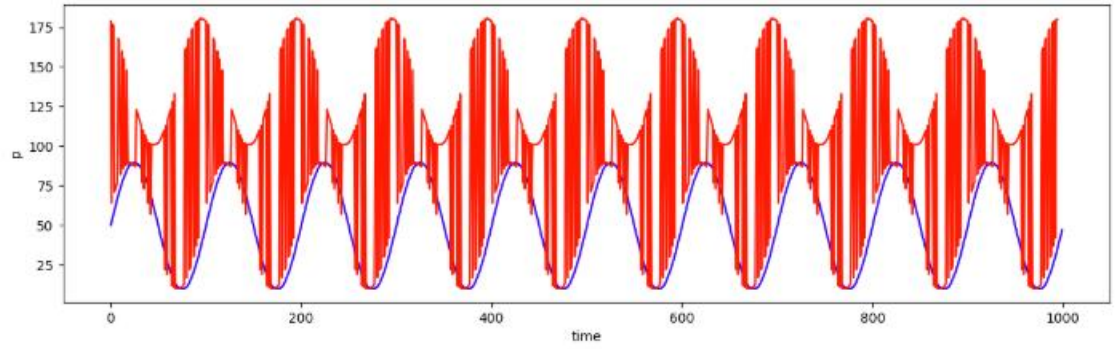
Question: In particular, compare the ML and MMSE estimates. What happened here?

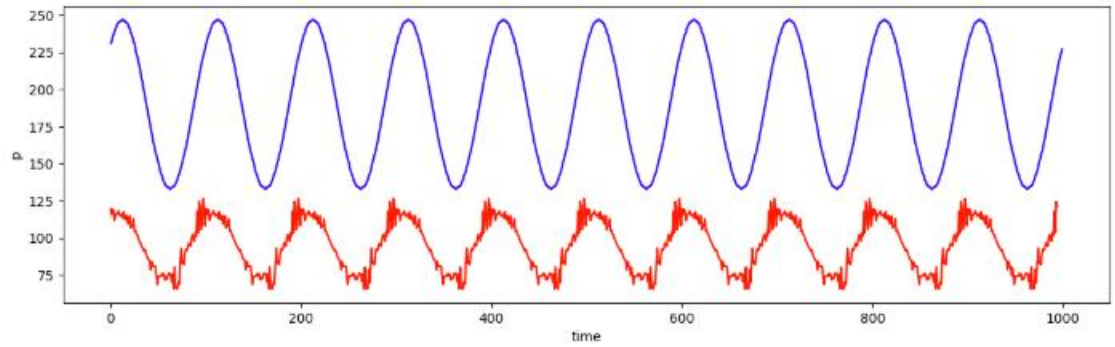
Answer: The ML estimate is much better than the MMSE estimate. We can see that both in the RMS error and the graphs. Thus, picking the winner based on the highest probability is much better than averaging (using MMSE). This makes sense as MMSE is better for numerical problems whereas ML is better for categorical problems, and this is a categorical problem as we are picking between a cosine and sine graph, we cannot have something in between (or averaged between the two).

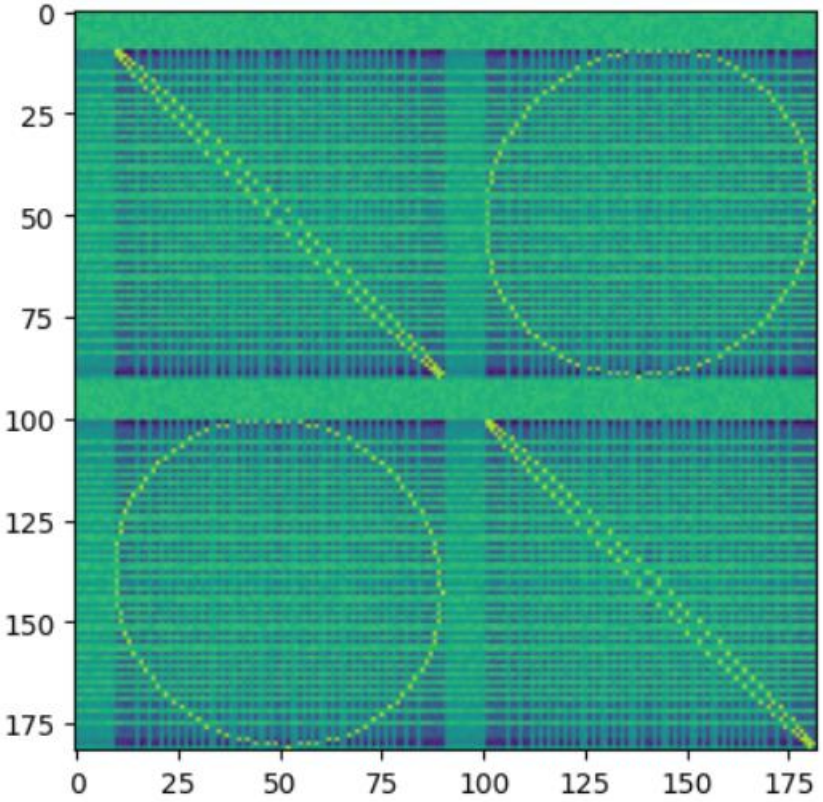
Question 3:

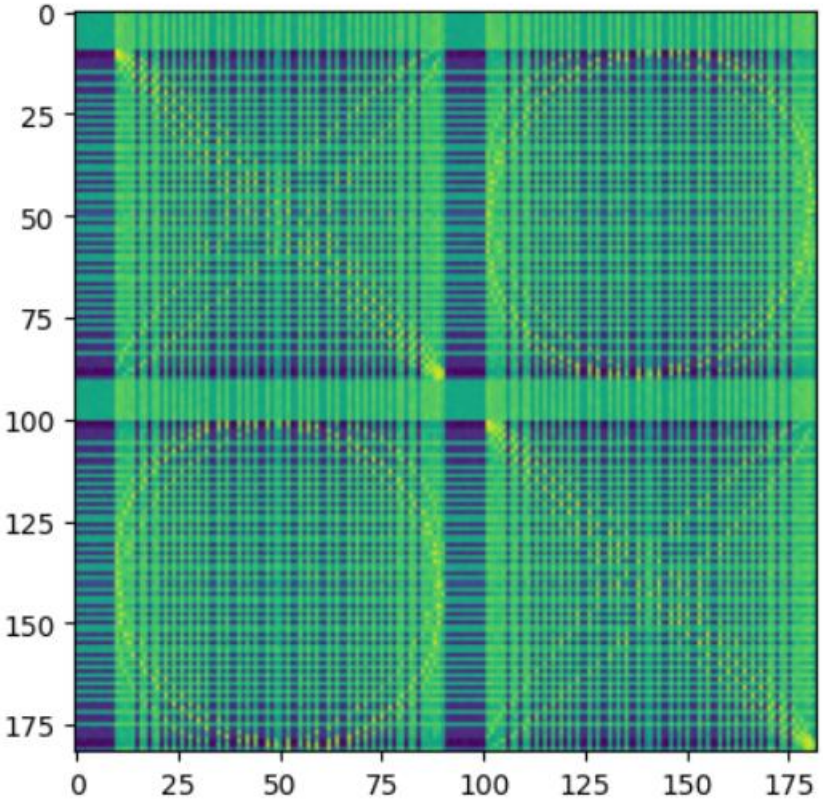



```
rms_error ML estimate 2285.338268178258
rms_error MMSE estimate 3102.400902182377
```






0
25
50
75
100
125
150
175
0
25
50
75
100
125
150
175


0
25
50
75
100
125
150
175
0
25
50
75
100
125
150
175

Question 3: Explain the weight matrices and ML and MMSE estimates. What is this a model of?

Both models do not do a good job and have substantial error. The ML estimate does a bit better. The weight matrices have clear defined patterns that they learned, neither is blurry, but as we can see these patterns do not get us a good estimate.