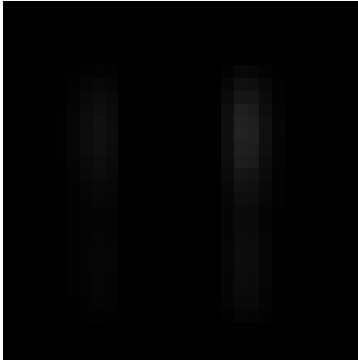


## MLSP HW3

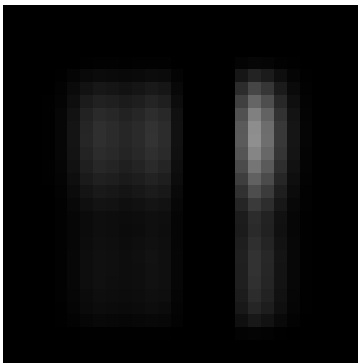
0880308 黃大祐

### 1. PCA

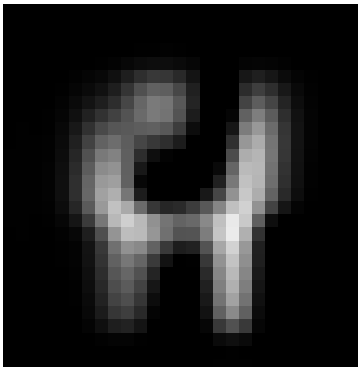
1X2



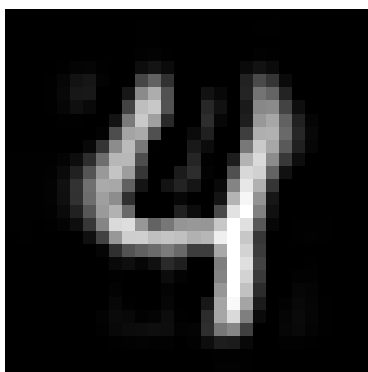
2X1



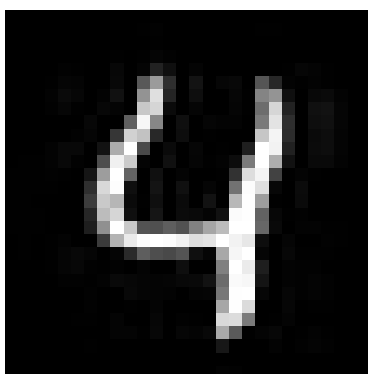
4X4



8X8



16X16



從上面可以看到，一開始只抓 2 個成分到 256 個成分，很明顯的圖片會越來越好，因為留下的資訊量越來越多，256 表示包含了 784 中資訊量最多的 256 個點，但相對的也會占比較大的空間。

## 2. ICA, NMF

在 ICA 及 NMF 的實驗中，我先分別實作了一次，接著再用 sklearn 的 FastICA 及 NMF 各做一次，令我感到十分驚訝的是 FastICA 的套件，裡面做的優化非常好，不過他的套件返回的資訊較少，像是這次要提取其中的 4 個主成分，就需要對原本的 function override。而我自己做的 ICA 由於沒有任何優化的部分(也沒用 FastICA 的演算法)，在計算上很慢，且效果不如預期，雖然可以看的出來分離出的主成分，但認為一定還有很多進步的空間。

而在 NMF，就是讓我最興奮的地方!!!!

只需要用梯度趨近最佳解，不停更新 W 跟 X 就能得到答案，計算上非常快，沒有用任何優化的方法也能有這樣的結果，真的讓我很佩服。

```
In [311]: print('It cost %f sec' % (tEnd - tStart))  
It cost 11.143527 sec
```

不僅如此，他分離的結果也非常好，1000 個 iteration 在 10 秒左右完成可以有這麼好的結果，而且未來在使用上也可以非常上手。

經過這次的作業，對於 **feature selection** 及 **dimension reduction** 多了非常多的理解，以前就只會用套件，也不知道裡面的數學怎麼推導，現在可以直接手推 **PCA** 真的讓我獲益良多！謝謝老師及助教的用心