EN.520.412.01. Machine Learning for Signal Processing
HW1
Youzhe Dou

Problem1:

1(b). The recomposed music has more noise than the original signal. This is because when performing pseudo inverse of a non-square matrix, the result is not accurate. When performing recomposing, the result will be noisier.

2(b). The recomposed music has more noise than the original signal.
It is slightly better than the recomposed music in 1(b).
In part 1, only individual notes is used in the projection and the results are combined afterwards. In part 2, all notes are used jointly and thus the result is better.

Problem2:

$$E = \frac{1}{DT}\|M - NW\|^2$$

$$\frac{dE}{dW} = \frac{1}{DT} * \frac{d\left((M - NW)(M - NW)^H\right)'}{dW} = \frac{2}{DT} * (-N)^H * (M - NW)$$

Problem3:

Gender detector using average face:

Dimension	50	100	200	300	500
Accuracy	56%	56.15%	56.2%	52.6%	52.6%

Gender detector using all faces in training set:

Dimension	50	100	200	300	500
Accuracy	56.8%	56.4%	56.7%	56.9%	56.9%

It can be observed that using all faces or using the average face does NOT make significant difference here. And using average face is MUCH faster.
This is why when using clustering, computing the average Euclidean distance from sample to all data in the cluster and only computing the distance between the sample and the average point of this cluster produce the same result and using average data is a much faster algorithm.