

CPSC 440/540 Machine Learning Assignment

Name:

Student ID:

1 Factor Analysis and PPCA [20 points]

- (1.1) [15 points] Implement PPCA algorithm with the closed-form solution in the space provided in `models.py`. Report the training and test error. (MLEs derived in the slides) (code). Also provide the reconstruction plot of test data provided when code is run.

Answer: Solution code provided with code package under solutions directory.

- (1.2) [2 points] Why do we use EM instead of closed form?

Answer: In space of high dimensionality, there may be computational advantage in using an iterative EM procedure rather than working directly with the sample covariance matrix.

- (1.3) [3 points] Both factor analysis and variational autoencoder are generative models, describe the similarities and differences of the two method.

Answer: They all assume that the latent space follows a Gaussian distribution. However, factor analysis uses linear projection to map latent representations to observed data, and variational autoencoder uses non-linear projection (MLP) to map latent representations to observed data.

2 Dimension Reduction Comparison [10 points]

Fashion MNIST is a dataset of images commonly used as a benchmark in computer vision research. It consists of 70,000 grayscale images of 28x28 pixels, each representing a clothing item from one of 10 categories, such as T-shirts, dresses, sneakers, sandals, and more. The dataset is intended to serve as a more challenging alternative to the classic MNIST dataset, which consists of handwritten digits.



Use the script provided, download the test set in Fashion Mnist, and generate the principle components (PCs) from PCA and factor analysis and independent components (ICs) from ICA. Change the random

state, do you see any differences? Describe the similarities of ICs and PCs you got, explain your observations, and submit the PCs and ICs at number of components is 10.

Answer: If we change the random state, we would observe that certain components flip the sign and the ordering of ICs changes. That is because of the ambiguity of these methods. In other word, flipping the sign won't change the value of likelihood of these methods, so sign of the components might change in different runs. FA and PCA give very similar results since they are both based on the assumption that $p(z)$ is Gaussian, but the order of PC4 and PC5 changes. ICA uses the non gaussian assumption so the results differ a lot. Even if ICA and other two methods uses different assumption, we can still see that certain components they find are similar. For example, IC8 is almost identical to PC(FA)4 and PC(PCA)5. PC1 and IC1 are quite similar-they both capture the shape of trouser.



