Linear Autoencoder ICP

David Sebesta

ECE 5258

Pattern Recognition

Fall 2021

Purpose

- Create a linear autoencoder
- Sample Covariance Matrix PCA
- Denoiseing

Theory

- Encoder
 - Sample Covariance Matrix PCA
 - Maps D to H dimensions
- Decoder
 - Maps H to D dimensions

Encoder

- Sample Mean
- Center Samples
- Sample Covariance Matrix
- EVD
- H Principal Components based on Total Variance Explained (p)

Encoding

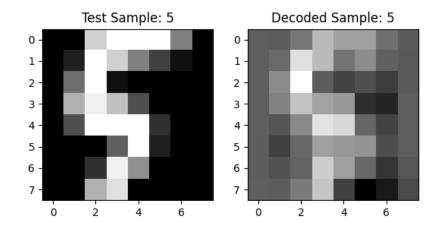
- Center test samples
- Transform with H Principal Components

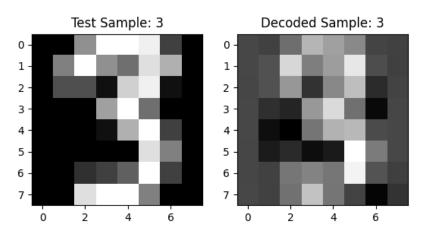
Decoding

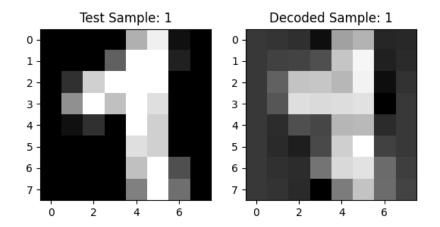
- Transform back with H Principal Components
- Decenter samples

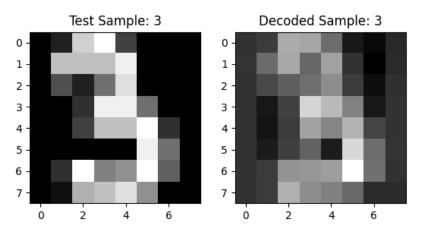
Data

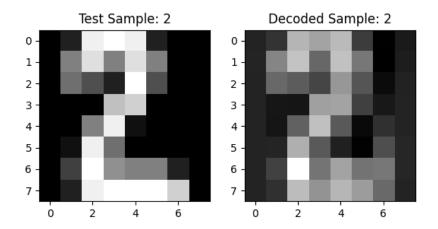
- Three Data Sets
 - MNIST 8x8 Digits
 - MNIST 28x28 Digits
 - Yale Faces
- Sample
 - D-length vector
 - Dimensions value = Pixel strength

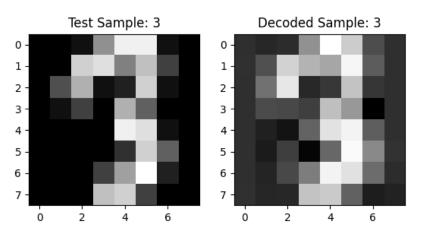


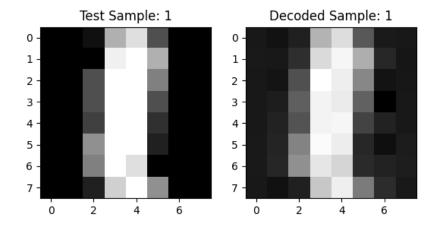


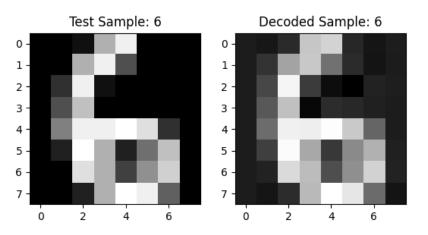


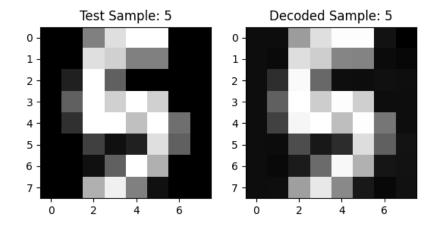


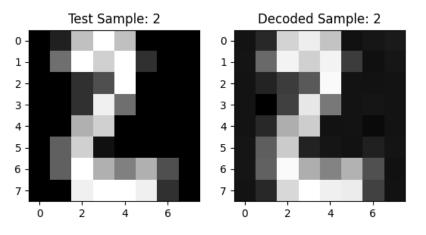






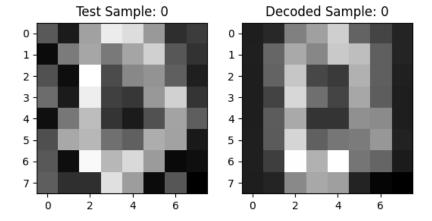




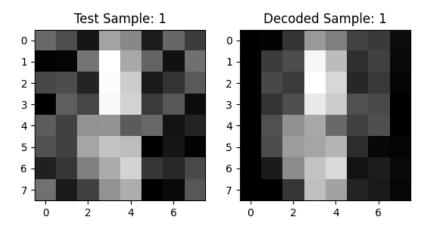


8x8 Digits with Noise



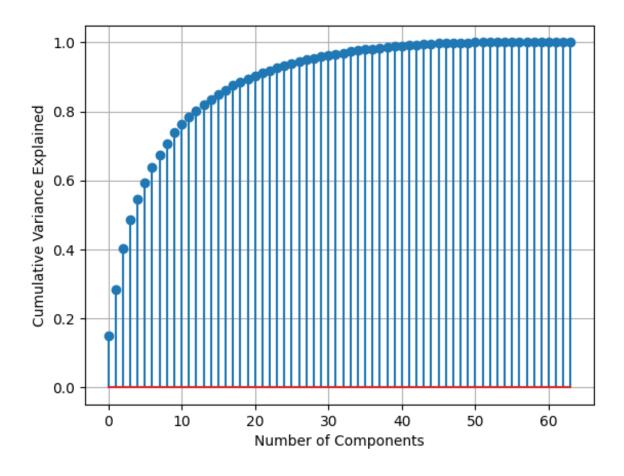


P = 0.95



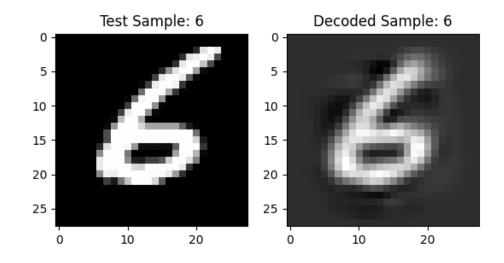
Cumulative Total Variance Explained

Majority in first few components



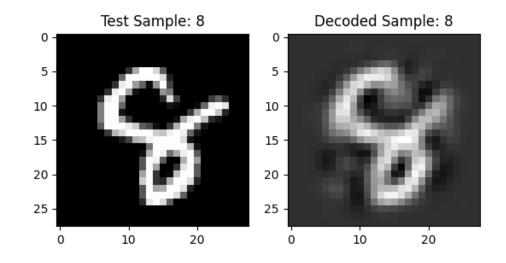
28x28 Digits

- P = 0.8
- H = 44



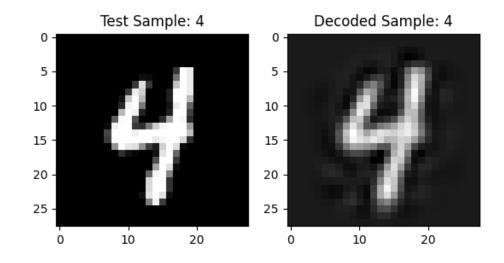
28x28 Digits

- P = 0.85
- H = 59

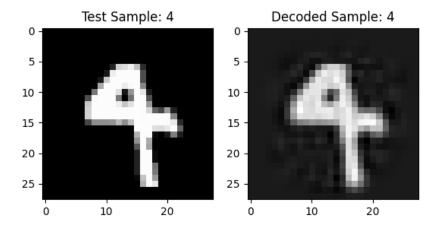


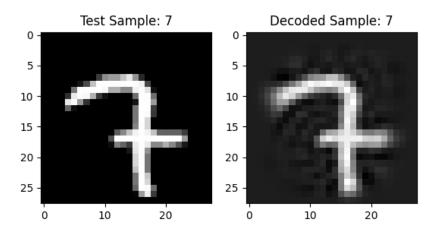
28x28 Digits

- P = 0.9
- H = 87

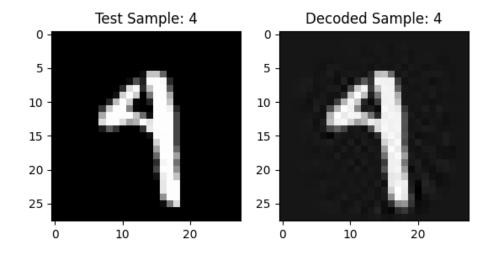


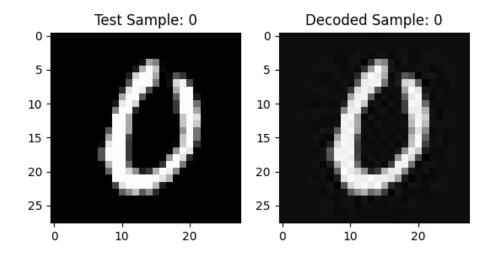
- P = 0.95
- H = 154

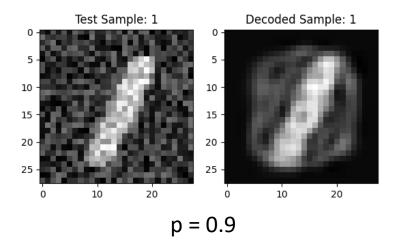


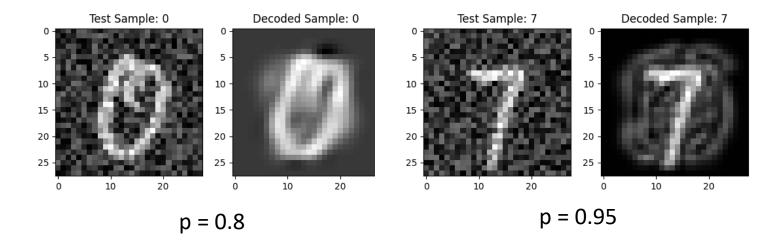


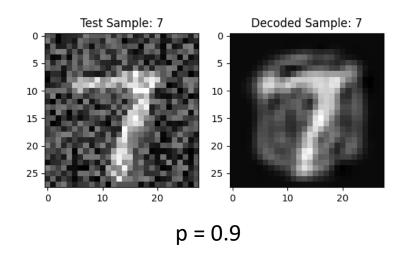
•
$$P = 0.99$$







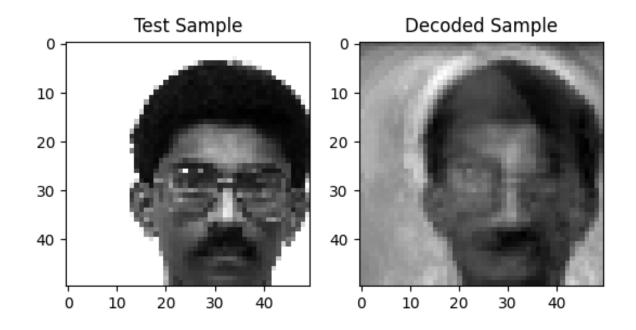




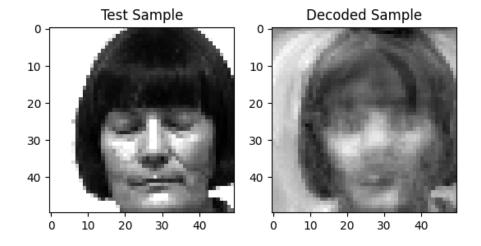
28x28 Digits with Noise

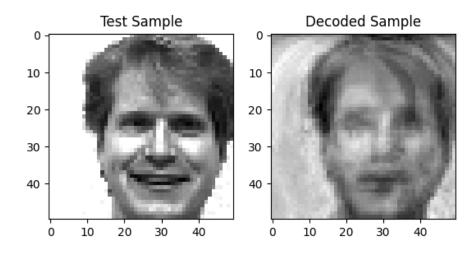
- 200 x 200
- Down sampled by 4
- D = 2500

- P = 0.8
- H = 11

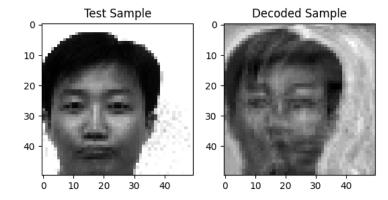


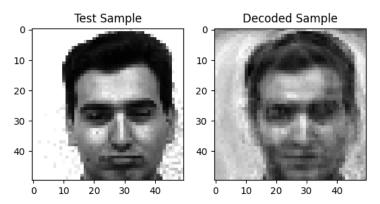
- P = 0.85
- H = 17

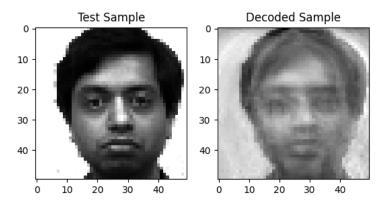




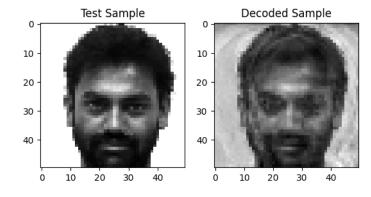
- P = 0.9
- H = 26

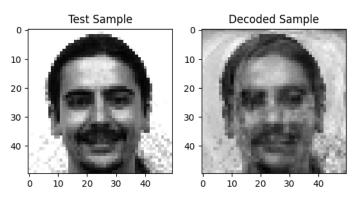


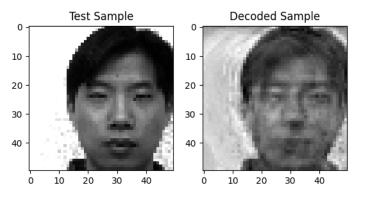




- P = 0.95
- H = 50







- P = 0.99
- H = 106

