

12-06_PCA

December 6, 2021

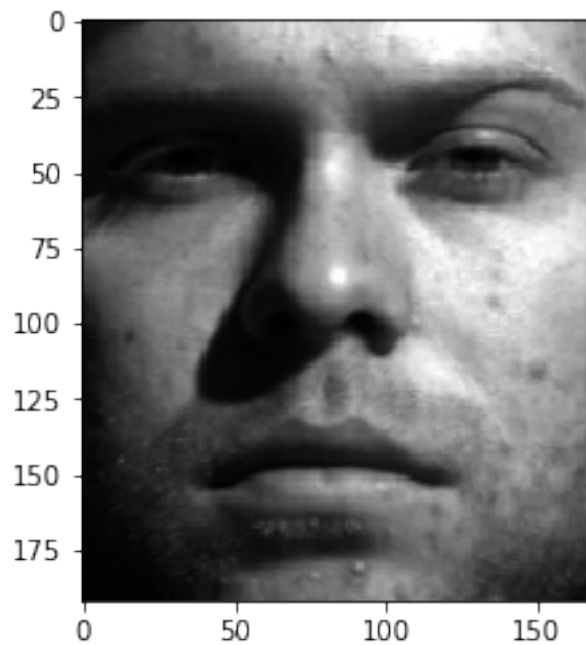
```
[1]: import numpy as np
import matplotlib.pyplot as plt
from PIL import Image
%matplotlib inline
```

The Cropped Yale dataset can be downloaded from [<http://vision.ucsd.edu/~iskwak/ExtYaleDatabase/ExtYaleB.h>]

```
[22]: import glob
face1 = []
for filename in glob.glob("CroppedYale/yaleB01/yaleB01_P00A*.pgm"):
    face1.append(np.array(Image.open(filename)))
face2 = []
for filename in glob.glob("CroppedYale/yaleB02/yaleB02_P00A*.pgm"):
    face2.append(np.array(Image.open(filename)))
```

```
[115]: plt.imshow(face1[0], cmap="gray")
```

```
[115]: <matplotlib.image.AxesImage at 0x7f27c8373b80>
```



```
[68]: X = np.hstack([face.reshape(-1,1) for face in face1[:-1] + face2[:-1]])
      X.shape
```

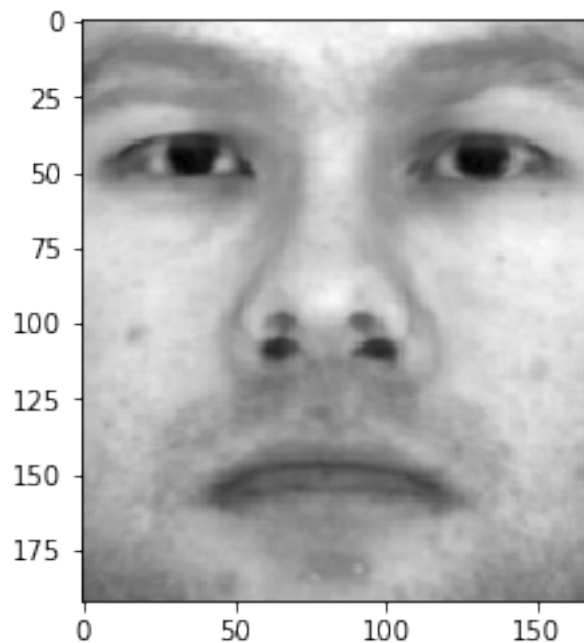
```
[68]: (32256, 126)
```

```
[69]: mean = np.mean(X, axis=1).reshape(-1,1)
      mean.shape
```

```
[69]: (32256, 1)
```

```
[70]: plt.imshow(mean.reshape(face1[0].shape), cmap="gray")
```

```
[70]: <matplotlib.image.AxesImage at 0x7f27f47be580>
```



```
[71]: U, S, V = np.linalg.svd(X - mean, False)
```

```
[72]: U
```

```
[72]: array([[ -1.52855097e-04,  -9.32562553e-03,   1.60595385e-03, ...,
           8.54073254e-04,  -2.70980743e-03,   1.59781071e-01],
         [-1.25621822e-04,  -9.27188471e-03,   1.73254178e-03, ...,
           6.95592547e-04,  -4.19604326e-03,   1.35114936e-01],
         [ 7.84827283e-05,  -9.42336112e-03,   2.19961771e-03, ...,
```

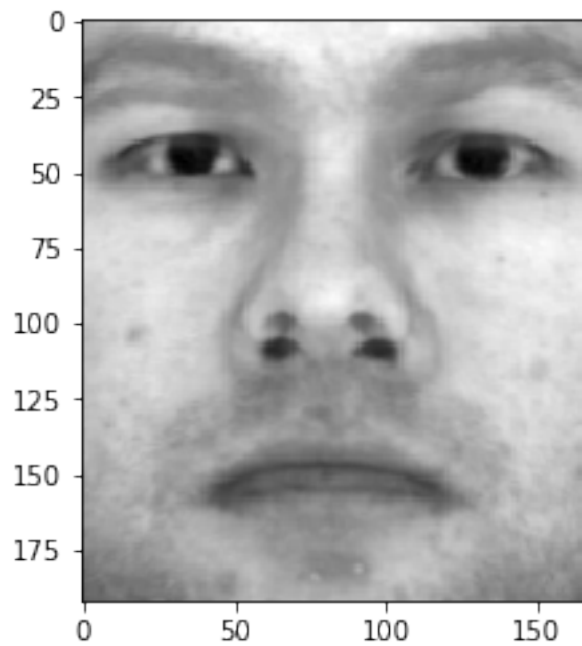
```

-7.92648516e-04, -5.65306736e-03, -6.80693860e-01],
...,
[ 1.07004969e-03,  5.35155374e-03, -9.69885794e-03, ...,
-3.82663533e-04,  3.78316296e-02, -1.48848558e-03],
[ 8.75517962e-04,  5.63841316e-03, -9.30946492e-03, ...,
-3.58117070e-04,  3.74531914e-02, -1.72597047e-03],
[ 7.00396114e-04,  6.03124220e-03, -8.84247489e-03, ...,
 2.23577066e-03,  4.20377713e-02, -2.28657180e-04]])

```

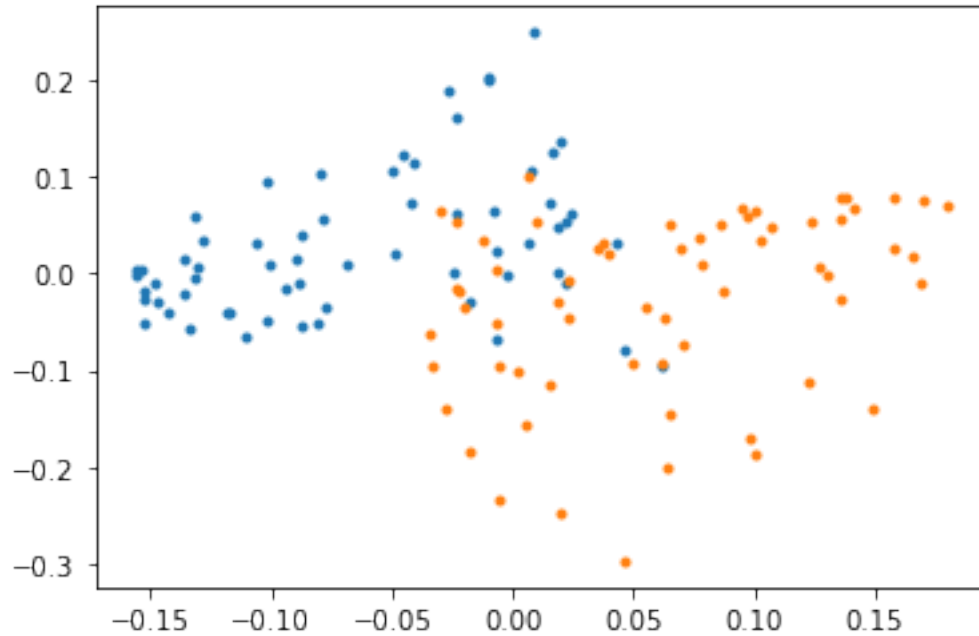
```
[73]: plt.imshow((U[:,0] + mean.flatten()).reshape(face1[0].shape) , cmap='gray')
```

```
[73]: <matplotlib.image.AxesImage at 0x7f27f470da90>
```



```
[94]: plt.plot(V[3,:63], V[9,:63], '.')
plt.plot(V[3,63:], V[9,63:], '.')
```

```
[94]: [<matplotlib.lines.Line2D at 0x7f27c87f4b20>]
```



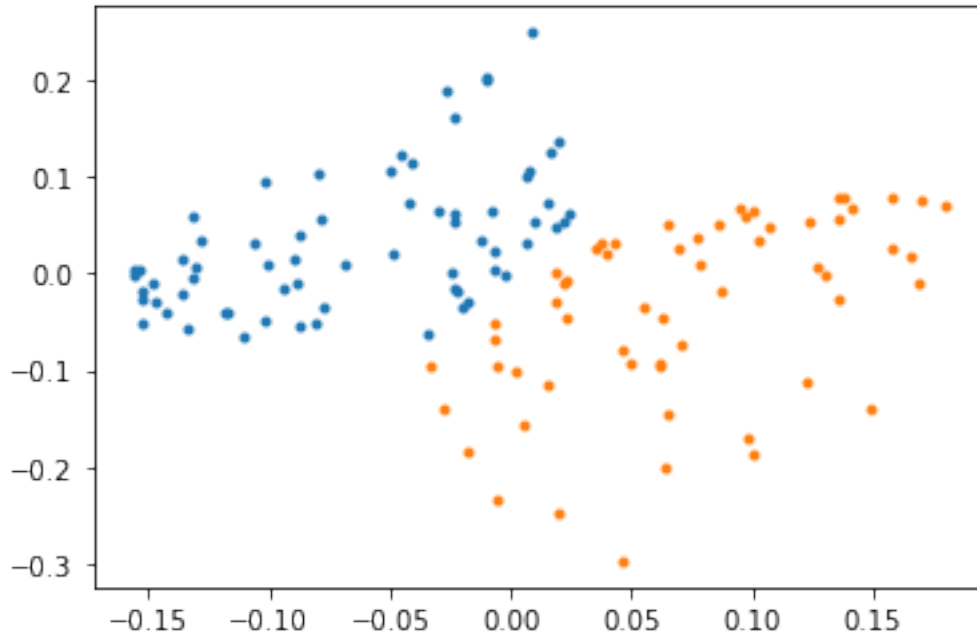
```
[78]: from sklearn.cluster import KMeans
```

```
[91]: cluster = KMeans(2).fit(V[(3,9),:].T)
```

```
[92]: pred = cluster.predict(V[(3,9),:].T)
```

```
[93]: plt.plot(V[3, pred==0], V[9, pred==0], '.')  
plt.plot(V[3, pred==1], V[9, pred==1], '.')
```

```
[93]: [<matplotlib.lines.Line2D at 0x7f27c8898fa0>]
```



```
[98]: (U.T @ U)[:3,:3]
```

```
[98]: array([[ 1.00000000e+00,  1.13949648e-15, -4.16333634e-17],
            [ 1.13949648e-15,  1.00000000e+00,  5.03937170e-16],
            [-4.16333634e-17,  5.03937170e-16,  1.00000000e+00]])
```

```
[99]: X_test = np.hstack((face1[-1].reshape(-1, 1), face2[-1].reshape(-1, 1)))
```

```
[101]: X_test.shape
```

```
[101]: (32256, 2)
```

```
[107]: V_test = (np.linalg.inv(np.diag(S)) @ U.T @ X_test)
V_test
```

```
[107]: array([[ 2.71978273e-01,  3.58788687e-02],
            [-3.60760290e-03,  4.86158849e-02],
            [ 5.88642592e-02,  1.00771768e-02],
            [-9.52715894e-02,  5.10484672e-02],
            [-2.06013194e-01, -1.77858915e-01],
            [-2.94850659e-02, -9.35870433e-03],
            [-6.11473717e-02,  1.29819824e-02],
            [-4.67701909e-04,  1.92202955e-01],
            [ 6.37512847e-03, -3.57244031e-02],
            [-1.13564494e-01, -2.14966827e-01],
            [ 3.16659074e-02,  8.47454304e-03],
```

[-2.57322198e-02, -5.95252018e-02],
[1.12009170e-01, 2.11698932e-02],
[-1.65784542e-01, -2.58917690e-01],
[1.27843782e-01, 7.00332556e-03],
[-5.81370403e-02, -2.66596111e-01],
[-1.90940507e-01, -2.41230111e-01],
[-1.88533457e-02, -9.49056763e-02],
[1.33595514e-01, 8.92773912e-02],
[-7.54862942e-02, -1.20013388e-02],
[-1.38400801e-01, -2.04967669e-01],
[-1.50165998e-01, -2.88029272e-01],
[1.04339490e-02, 1.92179561e-01],
[1.12603686e-01, 5.53919351e-02],
[3.34684304e-02, -8.56268293e-02],
[6.88262200e-02, 1.04135729e-01],
[-1.47937874e-01, -6.44437283e-02],
[9.41412656e-02, -2.70637753e-01],
[4.31926639e-02, -1.65139137e-01],
[-6.36458615e-02, -2.04838733e-01],
[2.16115840e-02, 8.40257425e-02],
[1.41989181e-01, 4.92700455e-02],
[-6.47749767e-02, -2.29568197e-01],
[5.29200360e-02, 6.05764727e-02],
[5.90801305e-02, 1.43185051e-01],
[3.30481752e-02, -1.38789495e-01],
[8.71553136e-03, -1.64793013e-01],
[-1.52550363e-01, -6.73117674e-02],
[4.51915166e-02, -2.39364803e-02],
[1.42079817e-02, 1.63448101e-01],
[-8.45311888e-02, -2.01021238e-01],
[-5.58335253e-02, 8.94611446e-02],
[-1.52265796e-02, -3.69126010e-02],
[2.36391044e-02, -1.18069143e-01],
[-5.91883252e-02, -4.78650404e-02],
[2.75692570e-03, -2.35427532e-01],
[3.31956235e-02, 7.59322470e-02],
[-2.07577896e-02, 2.13539950e-01],
[-7.42961782e-03, 4.07819584e-02],
[-1.01328717e-01, -1.04490864e-01],
[7.89041213e-03, 1.16188493e-02],
[1.44870904e-02, 9.27796555e-03],
[-9.41703127e-03, -1.80823187e-01],
[-3.94940958e-02, 2.10009104e-01],
[-1.33432462e-01, -1.31549207e-01],
[-1.06238602e-01, -1.17798801e-01],
[6.12722348e-02, 9.88403132e-02],
[-1.87463994e-02, 7.76999287e-02],

[6.66538314e-02, 3.03839594e-02],
[4.78143310e-02, 1.35277335e-01],
[1.76665420e-03, 1.12731605e-01],
[-6.80659856e-02, 4.32064099e-02],
[6.49306074e-02, -1.87562642e-02],
[-1.06762301e-01, -2.12731827e-01],
[-1.15269208e-01, -1.02280663e-01],
[1.67600071e-01, -4.39953809e-02],
[1.63637075e-01, 4.86936629e-02],
[3.07895934e-02, 2.29515055e-02],
[6.23119188e-02, 1.25486179e-02],
[3.63568226e-02, -7.60027747e-02],
[1.39022323e-02, 3.60680045e-02],
[-2.81958927e-02, 1.51518107e-01],
[-7.40880436e-02, -4.20728891e-02],
[-6.33995297e-02, 5.29862840e-02],
[-6.23360132e-02, 1.10023227e-01],
[-2.28017154e-02, -6.42440937e-02],
[8.55051451e-02, -1.50198983e-03],
[-1.20995505e-01, -1.30832505e-01],
[2.70137263e-02, -5.73894494e-02],
[7.44501405e-02, -6.47062524e-03],
[3.72932499e-02, -2.32823443e-02],
[-6.65751194e-02, -4.51141829e-02],
[1.42337402e-01, 2.12671415e-01],
[9.21655312e-02, 3.04939967e-02],
[1.11340779e-01, 5.68201910e-02],
[-1.42165745e-01, -2.84503102e-02],
[-6.04994785e-02, -1.19919613e-02],
[-8.61446598e-02, 8.94208344e-04],
[5.57066348e-02, -9.46039725e-03],
[-3.27701102e-02, 1.95217861e-01],
[-2.46506049e-03, -1.09872190e-01],
[2.01831983e-01, 3.08758384e-02],
[-7.75456101e-02, -1.61111012e-01],
[-9.49245565e-02, -2.94867458e-02],
[2.26936161e-01, 6.53925686e-02],
[-2.53986737e-01, 3.16693294e-02],
[-1.81814173e-01, 2.11767108e-02],
[-1.20423724e-01, 1.61225614e-01],
[1.33836894e-01, 2.74221063e-02],
[-5.03408431e-02, -7.67330994e-02],
[-2.28971077e-01, -1.25358301e-01],
[1.21407652e-01, 8.68394044e-02],
[2.81648667e-02, -1.60731216e-01],
[3.23426259e-02, 1.74316098e-01],
[-1.69762936e-02, 6.82867923e-02],

```

[ 2.56322344e-03, -2.80411035e-02],
[ 1.86820543e-01, -3.43345641e-02],
[-2.55730216e-01, -1.50606344e-01],
[ 7.56620690e-02,  2.25619857e-02],
[ 1.80376126e-01, -6.85734251e-02],
[ 4.16235083e-02, -9.34647306e-03],
[ 9.95502993e-02, -2.48798337e-01],
[ 1.15663419e-01,  1.60970302e-01],
[-3.82393022e-02, -2.47422856e-02],
[-2.10729710e-02,  2.20954310e-01],
[ 8.86727738e-02,  1.81954908e-01],
[-2.02498403e-01, -3.18220693e-02],
[ 1.61563018e-01,  4.58654666e-02],
[ 1.01136603e-01,  4.45516947e-01],
[-4.61953360e-02,  6.67931368e-05],
[-1.32591897e-01, -9.91700703e-03],
[-2.89009195e-01,  2.47080985e-01],
[-4.37272839e-02, -1.38758183e-01],
[-3.48887611e-02,  3.12584109e-02],
[ 1.50259372e-01,  5.47020788e-01],
[-9.52163526e+10, -3.70667670e+11]])

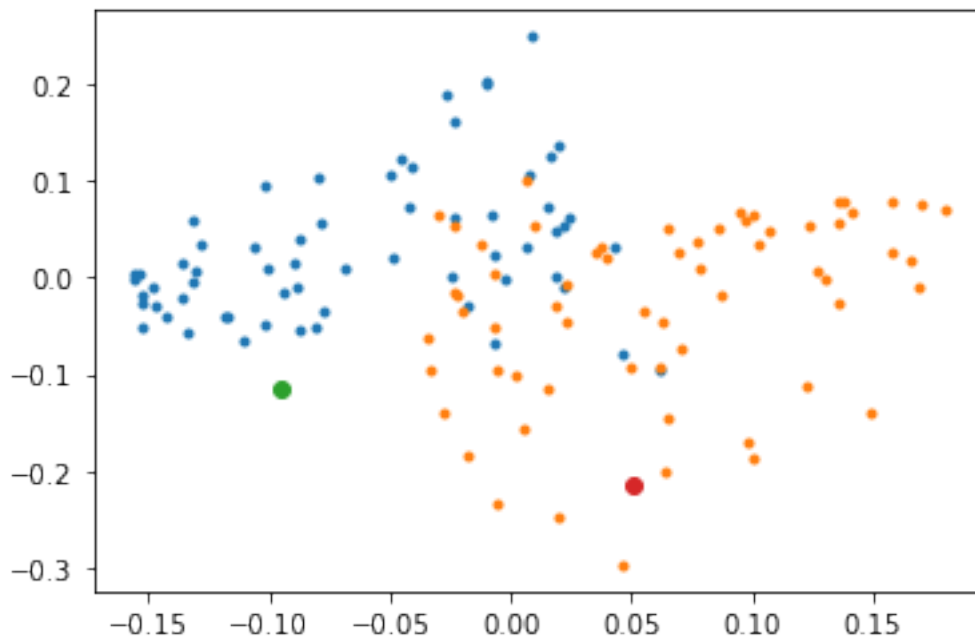
```

```

[112]: plt.plot(V[3,:63], V[9,:63], '.')
plt.plot(V[3,63:], V[9,63:], '.')
plt.plot(V_test[3,0], V_test[9,0], 'o')
plt.plot(V_test[3,1], V_test[9,1], 'o')

```

[112]: [<matplotlib.lines.Line2D at 0x7f27c8644be0>]

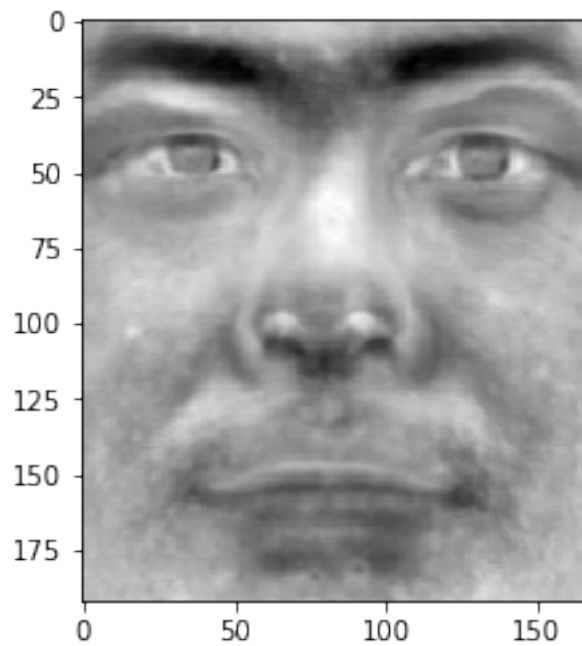



```
[113]: pred_test = cluster.predict(V_test[(3,9),:].T)
pred_test
```

```
[113]: array([0, 1], dtype=int32)
```

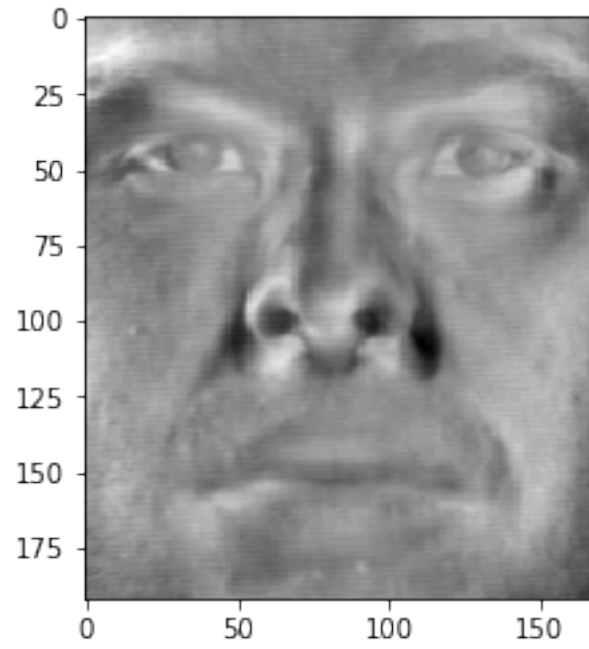
```
[116]: plt.imshow(U[:,3].reshape(face1[0].shape) , cmap='gray')
```

```
[116]: <matplotlib.image.AxesImage at 0x7f27c8358070>
```



```
[117]: plt.imshow(U[:,9].reshape(face1[0].shape) , cmap='gray')
```

```
[117]: <matplotlib.image.AxesImage at 0x7f27c82ac6d0>
```



[]: