Assignment 1: PCA

What you will learn

- Working with images using scikit-image
- PCA using scikit-learn
- Practical applications of PCA

Setup

- Download [Anaconda Python 3.6 (https://www.anaconda.com/download/)](https://www.anaconda.com/download/) for consistent environment.
- If you use pip environment then make sure your code is compatible with versions of libraries provided within Anaconda's Python 3.6 distribution.

Submission

- Do not change any variable/function names.
- Just add your own code and don't change existing code
- Save this file and rename it to be **studentid_lastname.ipynb** (student id (underscore) last name.ipynb) where your student id is all numbers.
- Export your .ipynb file to PDF (File > Download as > PDF via Latex). **Please don't leave this step for final minutes**.
- Submit both the notebook and PDF files.
- If you happen to use any external library not included in Anaconda (mention in **Submission Notes** section below)

Submission Notes

(Please write any notes here that you think I should know during marking)

[NO MARKS] PCA Warming Up (MUST READ)

I'm adding some code to illustrate examples of PCA using sklearn library.

Let's create some random `5d` data

In [67]:
```
import numpy as np
from sklearn.decomposition import PCA

# 100 points of 5d data
data = np.random.rand(100, 5)
```

Lets convert this `5d` data to `2d` using PCA`

In [68]:
```
# n_components=2 because I want to convert 5d data to 2d (dimensionality
 reduction)
pca = PCA(n_components=5)
pca.fit(data)
```

Out[68]:
```
PCA(copy=True, iterated_power='auto', n_components=5, random_state=Non
e,
  svd_solver='auto', tol=0.0, whiten=False)
```

In code above when we call `fit`, it populates two things in `pca`:

1. `mean_`
2. `components_`

In [69]:
```
# mean of the input data (per dimension) used to zeroying the mean
pca.mean_
```

Out[69]:
```
array([0.50122738, 0.51449788, 0.4896558 , 0.50737147, 0.52342601])
```

In [70]:
```
# basis vectors
pca.components_
```

Out[70]:
```
array([[-0.1188797 , -0.60421491, -0.33255902,  0.51015441,  0.4999389
2],
       [-0.60176066, -0.3100361 , -0.33962052, -0.64774673, -0.0827272
6],
       [ 0.61778037, -0.39141193, -0.43280807, -0.09339934, -0.5187464
2],
       [ 0.15894907,  0.57591333, -0.70809967, -0.09844896,  0.3632653
2],
       [-0.46565473,  0.23221216, -0.29212058,  0.54931868, -0.5849425
]])
```

Now we are ready to transform `5d` from `data` into `2d` using following code

In [71]:
```
data_to_reduce = data[:10]
reduced_data = np.dot(data_to_reduce - pca.mean_, pca.components_.T)

# reduced data from 5d to 2d
reduced_data.shape
```

Out[71]:
```
(10, 5)
```

You can accomplish the same using `transform` function provided in `pca`

In [72]:
```
pca.transform(data_to_reduce).shape
```
Out[72]:
```
(10, 5)
```

Time for inverse transform or changing `2d` data back to `5d`

Compression --> Decompression

In [73]:
```
decompressed_data = np.dot(reduced_data, pca.components_)+pca.mean_
decompressed_data
```
Out[73]:
```
array([[0.95099831, 0.82278359, 0.88030714, 0.23373863, 0.46764704],
       [0.79026944, 0.3796193 , 0.32334369, 0.86856097, 0.23011427],
       [0.68815806, 0.66955097, 0.5388003 , 0.46995356, 0.74164961],
       [0.45642941, 0.95794923, 0.04075203, 0.03749046, 0.82038421],
       [0.43090416, 0.64074393, 0.2412554 , 0.45572986, 0.12515528],
       [0.51447116, 0.66952889, 0.81283596, 0.67786059, 0.52856617],
       [0.74077576, 0.23177249, 0.44032238, 0.22887729, 0.08662266],
       [0.37543135, 0.07858668, 0.87394582, 0.09355736, 0.38358643],
       [0.69832257, 0.93892541, 0.98842144, 0.53449824, 0.06459605],
       [0.57492642, 0.47699855, 0.71957418, 0.7411886 , 0.46391866]])
```

In [74]:
```
# same can we accomplished using inverse_transform
pca.inverse_transform(pca.transform(data_to_reduce))
```
Out[74]:
```
array([[0.95099831, 0.82278359, 0.88030714, 0.23373863, 0.46764704],
       [0.79026944, 0.3796193 , 0.32334369, 0.86856097, 0.23011427],
       [0.68815806, 0.66955097, 0.5388003 , 0.46995356, 0.74164961],
       [0.45642941, 0.95794923, 0.04075203, 0.03749046, 0.82038421],
       [0.43090416, 0.64074393, 0.2412554 , 0.45572986, 0.12515528],
       [0.51447116, 0.66952889, 0.81283596, 0.67786059, 0.52856617],
       [0.74077576, 0.23177249, 0.44032238, 0.22887729, 0.08662266],
       [0.37543135, 0.07858668, 0.87394582, 0.09355736, 0.38358643],
       [0.69832257, 0.93892541, 0.98842144, 0.53449824, 0.06459605],
       [0.57492642, 0.47699855, 0.71957418, 0.7411886 , 0.46391866]])
```

In [75]:
```
# Lets find compression decompression error (absolute mean error)
np.sum(np.abs(data_to_reduce - decompressed_data))/data_to_reduce.size
```
Out[75]:
```
1.38222766565832e-16
```

Questions [8 marks]

Answer the questions below as follows:

1) What is 2+2

- 4
- 5
- 6

In [76]:
```
ans1 = 4
ans1
```

Out[76]: 4

2) (2 marks) For a n-D data, you can ALWAYS reconstruct the data with 0\% error if all `n` PCAs are used.

- True
- False

In [77]:
```
ans2 = True
ans2
```

Out[77]: True

3) (2 marks) From the 2nd tutorial, we ran PCA alorithm on faces. We called the extracted PCs--Eigenfaces. What is the value of a dot product between arbitary two eigen faces?

In [78]:
```
ans3 = ...
ans3
```

Out[78]: Ellipsis

4) (4 marks) Using the probablity, find the expected value for the function below:

Note: Lets say for a coin toss, head = 0 and tail = 1. Then the expected value for a coin-toss will be `p(x=tail)*1 + p(x=head)*0 = 1/2*1 = 0.5`.

In [79]:
```
def func():
    arr = [49, 8, 48, 15, 47, 4, 16, 23, 43, 44, 42, 45, 46]
    np.random.shuffle(arr)
    return min(arr[0:6])
```

Programming Tasks [92 marks]

Task 1: Building an Image Compression Algorithm

In this section you will build you own compression algorithm for images using PCA.

STEP 1: Read the image

1. Use `imread` function from `skimage.io` to read `leena.jpg`.
2. Show the image using `show_image` function provided to you

In [80]:
```
import matplotlib.pyplot as plt
# make matplotlib to show plots inline
%matplotlib inline

from skimage.io import imread

def show_image(img):
    if len(img.shape) == 2:
        plt.imshow(img, cmap='Greys_r')
    else:
        plt.imshow(img)
    plt.axis('off')
    plt.title('Image Dimension: {0},{1}'.format(img.shape[0], img.shape[
1]), fontsize=20)

# !!ADD CODE HERE!!
image = imread("leena.jpg")
#
#
show_image(image)
```

Image Dimension: 350,780

STEP 2: Compressing an image using PCA

Image is generally made of 3 (or 4) channels (RGB), we will build a compression algorithm that applies one channel at a time to compress multi-channel image.

In order to compress entire image you will compress all the channels within the image one by one and then serialize compressed channels and any auxiliary data required for decompression (for ex. principle components (components_), means (means_), original image size).

In order to decompress, you will deserialize the data, then uncompress the compressed channel one by one and stack them up to rebuild the uncompressed version of the original image.

Compression Strategy using PCA

- The above image you have read is of size `350 x 780` , i.e. width is `780` and height is `350`
- Patch the image into `10x10` patches yielding `35x78=2730` patches in total.
- Flatten each patch `(10x10)` into `100` dimensional vector.
- Now (for each channel) you will have `2730` number of `100-d` vectors.
- Apply PCA on these vectors.
- Reduce the dimensionality of `100-d` vector to `5-d` .

I've given you two functions `patchify` and `depatchify` .

`patchify` creates `100-d` vectors from all the patches from a given image and `depatchify` combines these `100-d` patches back to the image of the given size.

Please read through code below and figure out how these two functions work.

NOTE: `convert_to_cf` is important function to note (in cell below). It converts channel last format image to channel first format. Images that you read through `imread` function returns array of shape `X x Y x 3` channel is last axis. it is easier if channel were first axis then to extract anv channel vou can do use first indexer.

In [81]:
```
from sklearn.decomposition import PCA
import numpy as np


def patchify(img, ps=(10, 10)):
    patches = []
    h, w = img.shape
    for i in range(0, h, ps[0]):
        for j in range(0, w, ps[1]):
            patches.append(img[i:i+ps[0], j:j+ps[1]].ravel())
    return np.array(patches)


def depatchify(patches, patch_size=(10, 10), img_size=(350, 780)):
    # normalize
    patches[patches > 255.] = 255.
    patches[patches < 0.] = 0.

    # convert to unint8
    patches = patches.astype('uint8')
    rec_img = np.zeros(img_size, dtype='uint8')
    ph, pw = patch_size

    h, w = img_size
    x = 0
    for i in range(0, h, ph):
        for j in range(0, w, pw):
            rec_img[i:i+ph, j:j+pw] = patches[x].reshape((ph, pw))
            x += 1

    return rec_img

def convert_to_cf(img_cl):
    # convert image to channel first
    img_cf = np.swapaxes(img_cl.T, 1, 2)
    return img_cf

image=imread('leena.jpg')
img_cf = convert_to_cf(image)

# I'll patchify each channel and depatchify them
# Then I'll stack them together to create original image back
ch1 = depatchify(patchify(img_cf[0])) # first channel
ch2 = depatchify(patchify(img_cf[1])) # second channel
ch3 = depatchify(patchify(img_cf[2])) # third channel

# combine them now
rec_img = np.dstack((ch1, ch2, ch3))
plt.figure(figsize=(12, 12))
show_image(rec_img)
```

Image Dimension: 350,780

STEP 3: Compressing single channel of an image (FILL TWO FUNCTIONS BELOW)

Now you are familiar with how `patchify` and `depatchify` work. Do the following:

1. Write a function `compress` that will take one of the channel of the image as input and outputs compressed channel and auxiliary data required for decompressing. Return type of this function should be dictionary.
2. Write another function `decompress` that will take whatever dictionary data you returned from previous `compress` function and decompresses it into image channel (that was compressed).
3. PCA compression is lossy compression algorithm -- means you will loose the information during decompression.
4. I should be able to call two of your functions like so `decompress(compress(img_ch[0]))` to compress and decompress the given image's channel.

Compress: Pseudo code

- `Patchify` the given image's channel (input).
- Run `PCA` on the patches (you may use sklearn's implementation) to reduce them to `5d` vectors.
- You will need `basis vectors` or `principal components` and `mean` in order to reconstruct the data back to `100d`
- Return dictionary: `{'compressed_patches': 5d vectors, 'aux_data': principal components/basis vectors/means/final size of image}`
- By converting `100d` to `5d`, you reduce size by `20` times.

Decompress: Pseudo code

- Input is dictionary as returned by your `compress` function.
- Use `5d` vectors and do inverse PCA (check tutorial 2) and convert them back to `100d` vectors.
- Use `depatchify` function to convert these reconstructed `100d` vectors into an image channel

Tip: Good code is always modular and easy to read.

In [159]:
```
# COMPLETE FOLLOWING FUNCTIONS

def compress(img_ch, n_components=5):
    """
    Inputs
        img_ch: one of the channel of given image
        n_components: number of components returned by PCA
    Returns
        comp_data: Some data structure (may be dict.) that represents co
mpressed form of given input
        along with auxiliary data required for decompression (components
_, mean_ and shape of input image)
    """
    # patchify img_ch
    patches = patchify(img_ch)
    # ADD CODE HERE!!
    pca = PCA(n_components=n_components)
    pca.fit(patches)

    data = pca.transform(patches)

    result = {'y': data, 'aux_data': (pca.components_, pca.mean_, data.s
hape)}


    #
    #
    # return dictionary {'y': compressed data, 'aux_data': auxiliary dat
a for recontruction}
    return result

def decompress(comp_data):
    """
    Inputs
        comp_data: data structure that is returned by `compress` functio
n
    Returns
        img_ch: decompressed form of channel compressed and contained in
side `comp_data` data structure
    """
    #y = comp_data['y']
    #components, means, img_size = comp_data['aux_data']

    # recontruct the patches to 100d using aux_data and inverse PCA
    # depatchify and return img_ch
    # rec_patches = inverse_transform(y, components, means)
    # img_ch = depatchify(...)

    # create pca using aux data from comp_data
    pca = PCA()
    pca.components_ = comp_data['aux_data'][0]
    pca.mean_ = comp_data['aux_data'][1]
    data = pca.inverse_transform(comp_data['y'])

    img_ch = depatchify(data)
```

```
    return img_ch

# Red channel
# visualize your compression and decompression
img_ch = img_cf[0]
plt.figure(figsize=(10, 10))
show_image(img_ch)

plt.figure(figsize=(10, 10))
show_image(decompress(compress(img_ch, n_components=5)))
```

Image Dimension: 350,780

Image Dimension: 350,780

STEP 3: Compress and decompress entire image (FILL TWO MORE FUNCTIONS)

Write a `compress_image` function that:

- takes `channel last` (regular image read from `imread`) representation of an image
- convert `channel last` to `channel first` representation using `convert_to_cf` function defined previously
- compresses each of the channel using `compress` function
- outputs a one dictionary that contains all auxiliary data required to reconstruct/decompress entire image back.

Similarly, write `decompress_image` function that:

- decompresses the image compressed by `compress_image` function
- return `channel last` image

NOTE: You would patchify each channel and implement PCA on each channel thus for decompression you need `components_` , `mean_` for each channel and you need shape of input image as well.

In [85]:
```
def compress_image(img):
    """
    Inputs:
        img_cf: `Channel last` image data
    """
    img_cf = convert_to_cf(img)
    # compress each channel
    compressed_ch = []
    for channel in img_cf:
        compressed_ch.append(compress(channel))


    #create dictionary to return
    y_data = []
    aux_data = []
    for i in compressed_ch:
        y_data.append(i['y'])
        aux_data.append(i['aux_data'])

    result = {'y': y_data, 'aux_data': aux_data}

    return result

def decompress_image(comp_img):
    """
    Returns:
        img_rec: Decompressed `channel last` image
    """
    # gather channel data for image
    y0data = comp_img['y'][0]
    y1data = comp_img['y'][1]
    y2data = comp_img['y'][2]

    # gather channel aux_data for decompression
    aux0data = comp_img['aux_data'][0]
    aux1data = comp_img['aux_data'][1]
    aux2data = comp_img['aux_data'][2]

    # construct dictionaries for decompression input
    input0 = {'y': y0data, 'aux_data': aux0data}
    input1 = {'y': y1data, 'aux_data': aux1data}
    input2 = {'y': y2data, 'aux_data': aux2data}

    # gather decompressed image channels

    img_ch = np.dstack((decompress(input0), decompress(input1), decompre
ss(input2)))

    return img_ch

# test code
show_image(decompress_image(compress_image(imread("leena.jpg"))))
```

STEP 4: Serialization and de-serialization

You can easily (de)serialize dictionary using `np.save` and `np.load` (see example below)

`compress_and_serialize` should:

- Read image given by `inp_path`
- Use `compress_image` to compress it
- Serialize the compressed data using `np.save` to a file specified by `out_path`

NOTE: All the data required for deserialization must be saved to a one single file only.

`deserialize_and_decompress` should:

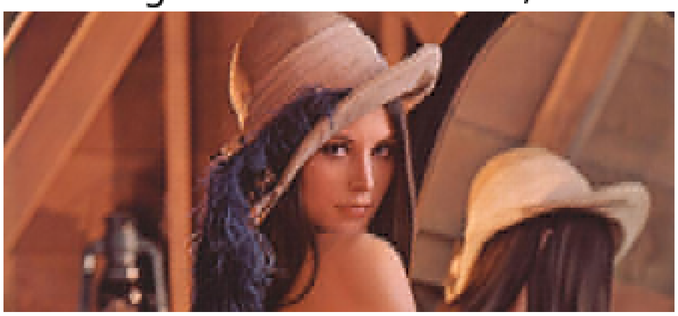
- Read the file specified by `inp_path` using `np.load`
- Use `decompress_image` function to decompress it
- Return image (channel last as returned by `decompress_image`) function

In [84]:
```
# EXAMPLE OF SAVING AND LOADING DICTIONARY OBJECT TO/FROM FILESYSTEM
d = {'a': [1, 2, 3], 'b': [4, 5, 6, 7, 8]}
np.save('example.npy', d)
ds = np.load('example.npy').item()

ds
```

Out[84]:
```
{'a': [1, 2, 3], 'b': [4, 5, 6, 7, 8]}
```

In [105]:
```
# COMPLETE THESE TWO FUNCTIONS
def compress_and_serialize(inp_path='leena.jpg', out_path='output.bin'):

    img = imread(inp_path)

    # compress the image
    compressed_img = compress_image(img)
    y = compressed_img['y']
    aux_data = compressed_img['aux_data']
    # serialize image to output path
    np.savez(out_path, y=y, aux_data=aux_data)

def deserialize_and_decompress(inp_path='output.bin.npz'):
    serialized_img = np.load(inp_path)
    decompressed_img = decompress_image(serialized_img)

    return decompressed_img


compress_and_serialize()
show_image(deserialize_and_decompress())
```

Image Dimension: 350,780



STEP 5: What is size of output.bin file?

Did you end of in compressing anything? Conclude your experiments!

Conclusion

If we compare the size of the image vs size of the compressed data outputted by the compress_image() function, we can tell that it is significantly smaller in size than the original image. The original image is

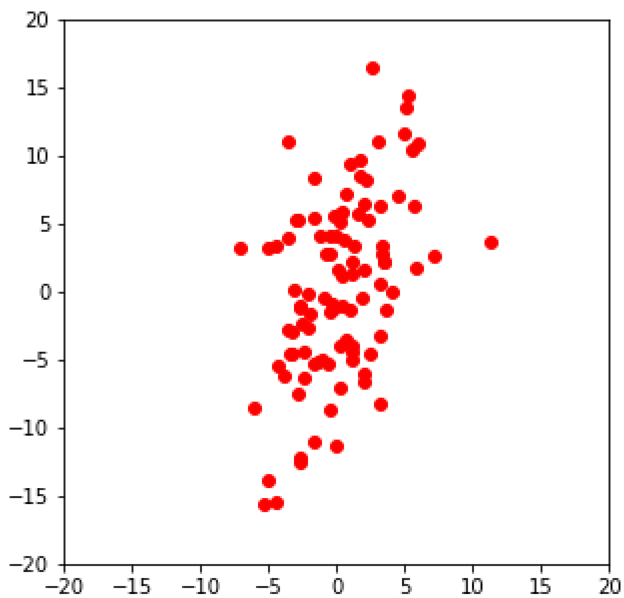
In [104]:
```
import os
import sys

print('Original', os.path.getsize('leena.jpg'), 'bytes')
print('Compressed', os.path.getsize('output.bin.npz'), 'bytes')

img = imread('leena.jpg')

# compress the image
compressed_img = compress_image(img)
print(sys.getsizeof(compressed_img))
print(sys.getsizeof(img))
```

```
Original 35740 bytes
Compressed 350458 bytes
240
819128
```

Task 2: Rotation and Translation Invariance in PCA

STEP 1 (done already): Create normally distributed data

In [142]:
```
mean = [0, 0]
cov = [[10, 10], [10, 40]]  # diagonal covariance
x, y = np.random.multivariate_normal(mean, cov, 100).T
X = np.array(list(zip(x, y)))

def plot_data(X):
    plt.figure(figsize=(5, 5))
    plt.plot(X[:, 0], X[:, 1], 'ro')
    plt.xlim([-20, 20])
    plt.ylim([-20, 20])
    plt.gca().set_aspect('equal', adjustable='box')
    plt.show()

plot_data(X)
```

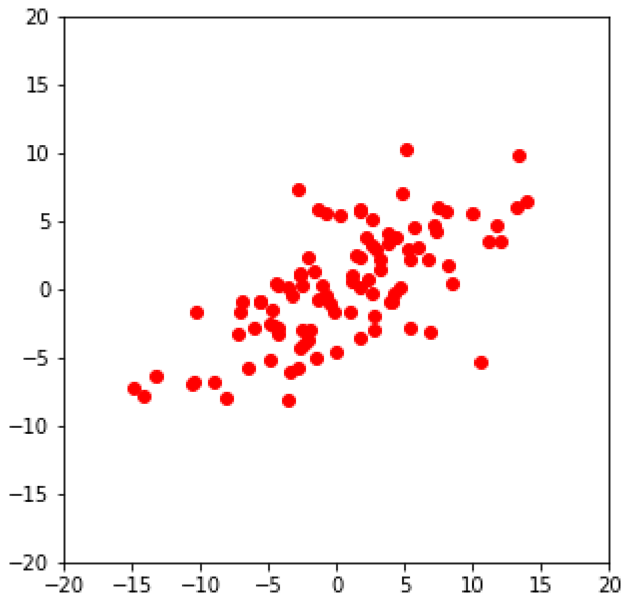


STEP 2: Rotate the data by 45 degrees and create new array X_rot

Write code to rotate X by 45

- Use rotation matrix
- Or individually rotate each point in `X` by 45 degrees
- Use `plot_data` function defined in cell above to visualize the rotated data

In [143]:
```
theta = np.radians(45)
# WRITE CODE HERE
# Code to rotate X
c, s = np.cos(theta), np.sin(theta)
R = np.array(((c,-s), (s, c)))
# print(X)
X_rot = np.matmul(X, R)
# populate new matrix X_rot
plot_data(X_rot)
```



STEP 3 (done but need explanation): Perform pca with n_components=2

Do you see anything interesting?

Explain the code below and write down your observations (2-3 lines only).

Observing the chart of data, we can see that the data produced by the pca for both the regular X and X_rot gave the exact same values.

Observations

WRITE YOUR ANSWER HERE

In [144]:
```
import pandas as pd

def visualize_components(X, X_rot):
    pca = PCA(n_components=2)

    pca.fit(X)
    df1 = pd.DataFrame(pca.fit_transform(X))
    df2 = pd.DataFrame(pca.fit_transform(X_rot))

    return pd.concat((df1, df2), axis=1)[:20]

visualize_components(X, X_rot)
```

Out[144]:

	0	1	0	1
0	-9.412424	-0.955884	-9.412424	-0.955884
1	-1.608987	-5.281836	-1.608987	-5.281836
2	-8.918562	-1.546539	-8.918562	-1.546539
3	-7.061726	3.783412	-7.061726	3.783412
4	-0.540785	3.847406	-0.540785	3.847406
5	-1.039686	0.694933	-1.039686	0.694933
6	1.163761	-3.162520	1.163761	-3.162520
7	-8.039922	-0.120444	-8.039922	-0.120444
8	-3.666930	2.235901	-3.666930	2.235901
9	3.670908	1.459503	3.670908	1.459503
10	-3.383177	-0.462185	-3.383177	-0.462185
11	-14.832154	1.286534	-14.832154	1.286534
12	3.547446	-1.345116	3.547446	-1.345116
13	-6.255115	0.210463	-6.255115	0.210463
14	7.148332	-0.712905	7.148332	-0.712905
15	6.427158	3.566705	6.427158	3.566705
16	-2.789798	5.093354	-2.789798	5.093354
17	-3.926460	6.119776	-3.926460	6.119776
18	4.109782	-2.421226	4.109782	-2.421226
19	-11.656337	2.841705	-11.656337	2.841705

STEP 4: Perform PCA again and find angle between principal components

- Perform PCA on X and X_rot with n_components=1
- It will give one basis vector for each of data (X and X_rot)
- Find the angle between these two basis vectors
- Explain your observations?

Observations

WRITE YOUR OBSERVATIONS HERE

In [147]:
```
from numpy.linalg import norm

def angle_between(a,b):
    # COMPLETE THIS FUNCTION
    # CALCULATE ANGLE IN RAD BETWEEN TWO VECTORS
    unit_a = a / np.linalg.norm(a)
    unit_b = b / np.linalg.norm(b)
    return np.arccos(np.clip(np.dot(unit_a, unit_b), -1.0, 1.0))

pca = PCA(n_components=1)
pca.fit(X)
c1 = pca.components_

pca.fit(X_rot)
c2 = pca.components_

np.rad2deg(angle_between(c1[0], c2[0]))
```
Out[147]:
```
44.999999999999986
```

(NO MARKS) STEP 5: Repeat these experiments for translation as well

There is not marks for this part. You can do this for your own learning.

Now translate every point in `X` by fixed `x` and `y` amount.

```
X = X + [1, 2]
```

like so and repeat all the above experiments in cell below and write down your observations:

In [782]:
```
# PERFORM EXPERIMENTS WITH TRANSLATION HERE
# NO MARKS FOR DOING THIS
```

TASK 3: Recovery of corrupted images using PCA

Check the code below.

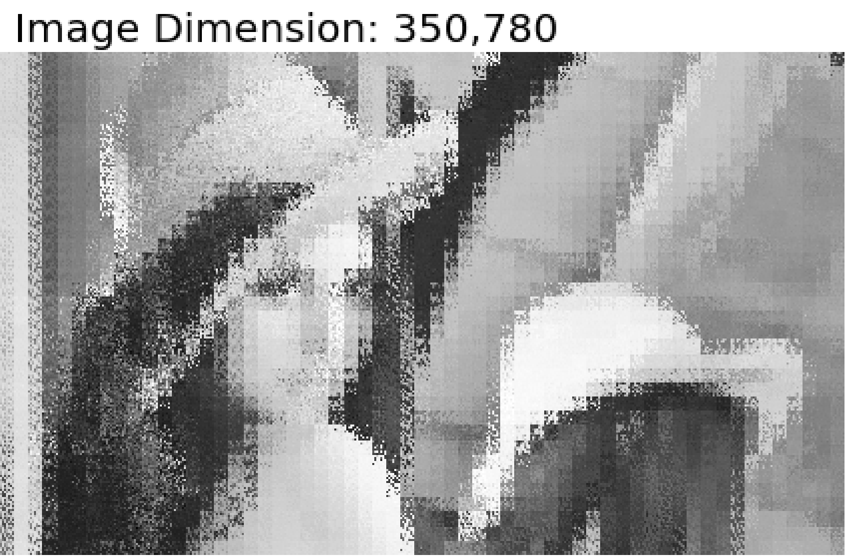
It corrupts the `leena.jpg` image that you worked on before.

Check very carefully what below code is doing on the image and answer:

Is it same as rotation of data points (like done before), if yes explain (just one liner)?

EXPLAIN HERE

In [148]:
```
# Code for corrupting the leena image
image = convert_to_cf(imread('leena.jpg'))[0]

def corrupt_image(img):
    # lets patchify R channel with patches of 35x78 patches
    patches = patchify(img)

    # Noise 1:
    # lets jumble pixels of patches now
    jumble_idx = list(range(len(patches[0])))
    np.random.shuffle(jumble_idx)

    jumbled_patches = np.array([patch[jumble_idx] for patch in patches])
    rec_jumbled_image = depatchify(jumbled_patches, img_size=img.shape)

    return rec_jumbled_image


cimage = corrupt_image(image)

plt.figure(figsize=(10,10))
show_image(image)

plt.figure(figsize=(10,10))
show_image(cimage)
```



Recovering from the corruption

Use the `compress` function you coded earlier to get compressed form of original leena image and corrupted leena image.

We will try to recover the corrupted leena given the compressed version of original leena.

Notice in code below, I replace `aux_data` of corrupted version with `aux_data` of original version.

Does the code below code work in recovering the original leena image back?

Explain how below code works and show it actually works?

EXPLAIN HERE

In [150]:
```
# This assuming your compress function returns dictionary which contains
 `aux_data`
# if you coded your compress in diferent way then change code below appr
opriately
# basically u need to replace auxiliary data of corrupted with original
 while keeping compressed (transformed) data same
d = compress(image, n_components=50)
d1 = compress(cimage, n_components=50)

plt.figure(figsize=(10, 10))
show_image(decompress(d1))

d1['aux_data'] = d['aux_data']
plt.figure(figsize=(10, 10))
show_image(decompress(d1))
```

Image Dimension: 350,780

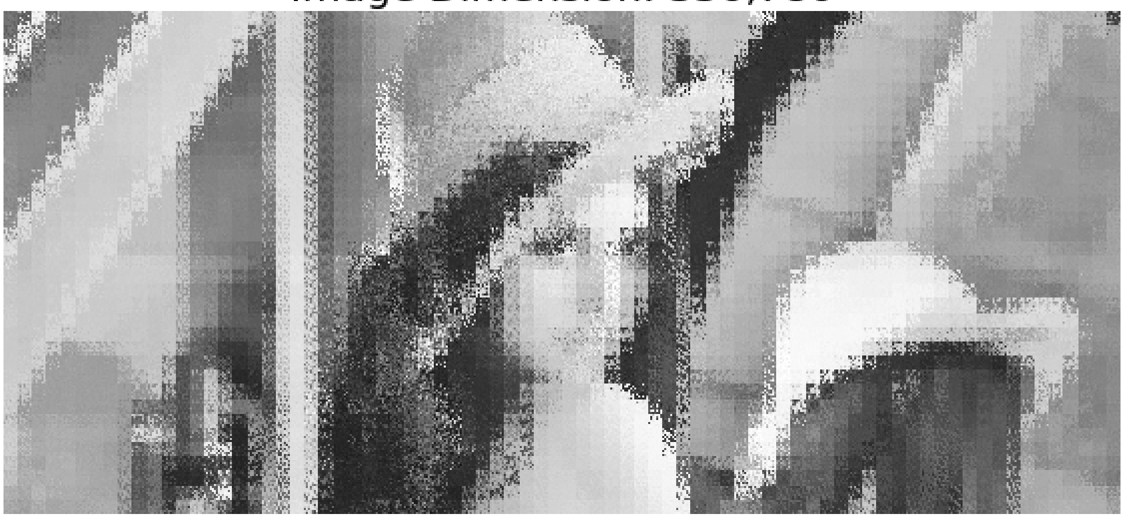


Image Dimension: 350,780

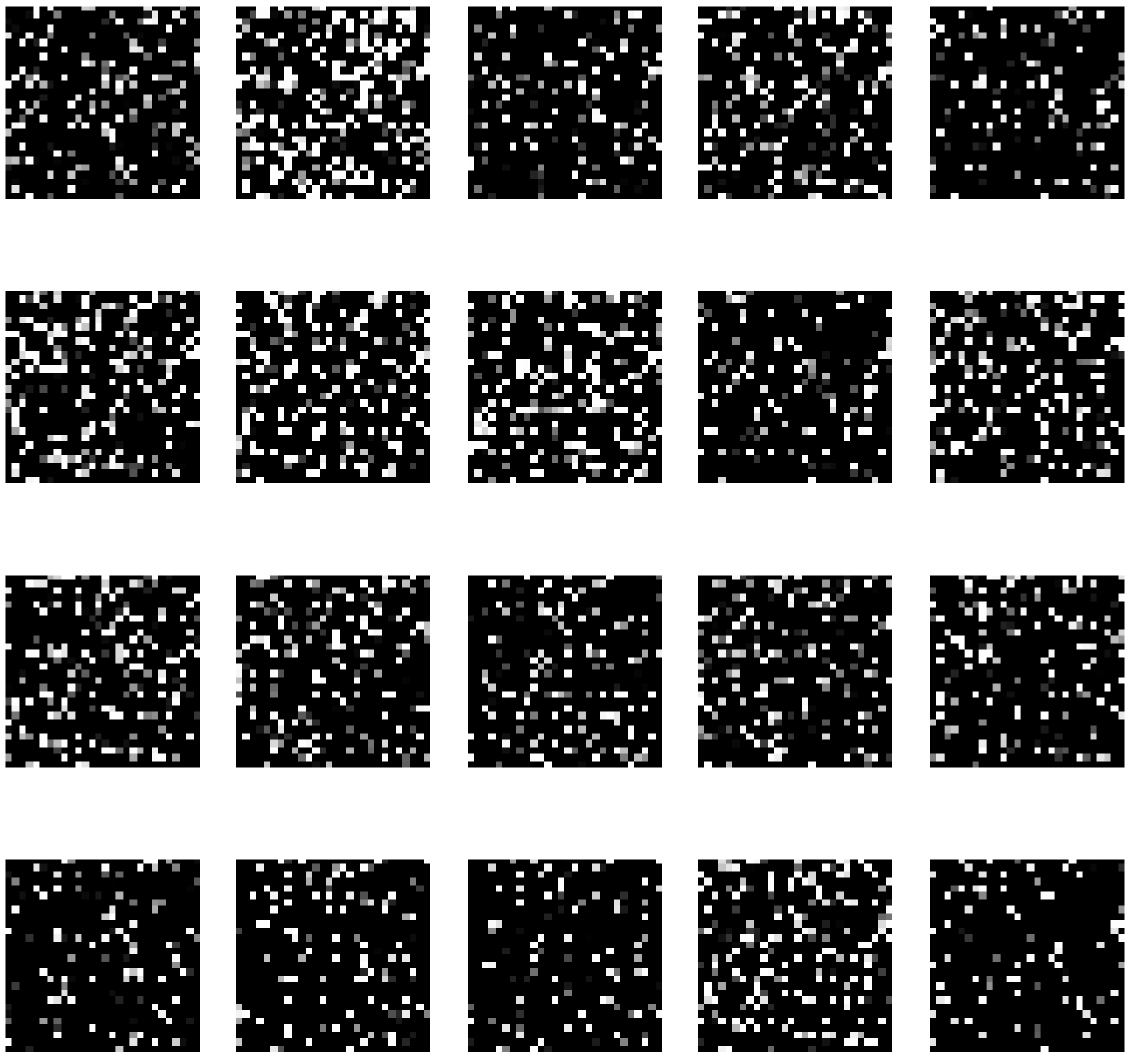
Task 4: Decrypting manuscript of the lost civilization

You belong to one of the advanced civilization, while exploring the universe you land on the planet Earth. However, there is no inhabitants on the planet anymore. While exploring Earth, you come across some damaged hard drive that contains various images. You suspect that these images form a manuscript of how "humans" use to write different digits in maths. You recovered all the data from the hard drive safely. You opened up your jupyter notebook (python being universal language and popular among alien species), you started plotting the images you just recovered.

In [161]:
```
rimages = np.load('recovered_images.npy')
rimages.shape
```

Out[161]:
```
(20, 28, 28)
```

In [152]:
```
def plot_manuscript(images):
    for i, img in enumerate(images):
        plt.subplot(4, 5, i+1)
        plt.imshow(img, cmap='Greys_r')
        plt.axis('off')

plt.figure(figsize=(10, 10))
plot_manuscript(rimages)
```



You being a smart alient, figured out that these images are encrypted and not in their original form. You also figured out that encryption is just fixed jumbling of the pixels of the images. Since you're unaware of the manuscript, you cannot decrypt the images just by themselves even if they follow the fixed jumbling pattern. However, fortunately you found the `principal components` and `mean` of the the original manuscript somewhere in the same harddisk. Now, your task is to recover all the 20 images and plot them nicely in cell below.

Use `plot_manuscript` function from above to plot the recovered manuscript.

In [166]:
```
original_components = np.load('original_pca.npy')
original_mean = np.load('original_mean.npy')
print(original_components.shape)
# Write your code here
pca = PCA(n_components=3)

pca.mean_ = original_mean
pca.components_ = original_components
print(rimages)
pca.inverse_transform(rimages)
# plot manuscript
```

```
(20, 784)
[[[  0   0   0 ... 108   0   0]
  [  0   0   0 ...   0   0  93]
  [  8   0   0 ... 199   0   0]
  ...
  [  0   7   0 ... 234   0   0]
  [  0 210   0 ...   0   0   0]
  [  0   0   0 ...   0   0   0]]

 [[  0   0   0 ... 132   0   0]
  [  0   0  59 ... 252   0   0]
  [252  67   0 ... 253   0   0]
  ...
  [  0 252   0 ... 217   0   0]
  [  0 252   0 ...  52   0   0]
  [  0   0   0 ...   0   0   0]]

 [[  0   0   0 ...   0   0   0]
  [  0   0   0 ...   0   0   0]
  [  0   0   0 ...   0   0   0]
  ...
  [  0   0   0 ...  40   0   0]
  [  0  98   0 ...   0   0   0]
  [  0   0   0 ...   0   0   0]]

 ...

 [[  0   0   0 ...   0   0 253]
  [  0   0   0 ...   0   0   0]
  [  0   0   0 ...   0   0   0]
  ...
  [  0   0   0 ...  60   0   0]
  [ 23   0   0 ... 253   0   0]
  [  0   0   0 ...   0   0   0]]

 [[  0   0   0 ...   0   0   0]
  [  0   0   0 ...   0   0   0]
  [252   0   0 ... 252   0   0]
  ...
  [  0   0   0 ...   0   0   0]
  [ 32   0   0 ... 252   0   0]
  [  0   0   0 ...   0   0   0]]

 [[  0   0   0 ...   0   0   0]
  [  0   0   0 ...   0   0  37]
  [207   0   0 ...   0   0   0]
  ...
  [  0   0   0 ...   0   0   0]
  [207   0   0 ...   0   0   0]
  [  0   0   0 ...   0   0   0]]]
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
<ipython-input-166-eb86816b1562> in <module>
      8 pca.components_ = original_components
      9 print(rimages)
---> 10 pca.inverse_transform(rimages)
     11 # plot manuscript

/anaconda3/lib/python3.7/site-packages/sklearn/decomposition/base.py in inverse_transform(self, X)
    158                             self.components_) + self.mean_
    159         else:
--> 160             return np.dot(X, self.components_) + self.mean_

ValueError: shapes (20,28,28) and (20,784) not aligned: 28 (dim 2) != 20 (dim 0)
```

```
In [ ]:
```