Student Names and IDs:Yiteng Lu(0882828);Wenge Xie(0856177);Zengtian Deng(0877187)

- Student 1, ID1 (Replace this item with your first and last name and student ID number. Add more items like this as needed, including the > - characters at the beginning, which generate the indent and the bullet.)

Homework 1

Part 1: Calculus

Problem 1.1 ¶

Write an expression for the gradient of f,

$$\nabla f(\mathbf{x}) = \frac{\partial f}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} .$$

Answer

$$\nabla f(\mathbf{x}) = \frac{\partial f}{\partial \mathbf{x}} = \begin{bmatrix} 3x^2 - 6x \\ 3y^2 - 6y \end{bmatrix}$$

Problem 1.2

Write an expression for the Hessian of f,

$$H_f(\mathbf{x}) = \frac{\partial^2 f}{\partial \mathbf{x} \partial \mathbf{x}^T} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}.$$

Answer

$$H_f(\mathbf{x}) = \frac{\partial^2 f}{\partial \mathbf{x} \partial \mathbf{x}^T} = \begin{bmatrix} 6x - 6 & 0 \\ 0 & 6y - 6 \end{bmatrix}$$

Problem 1.3

Give expressions for the eigenvalues $\lambda_1(x, y)$ and $\lambda_2(x, y)$ (in any order) of H_f.

Answer

$$\lambda_1(x, y) = 6x - 6$$
$$\lambda_2(x, y) = 6y - 6$$

Problem 1.4

Give expressions for all the unit-norm eigenvectors $\mathbf{u}_1$ of H_f corresponding to the eigenvalue $\lambda_1(x, y)$ you found in the previous problem and for all the unit-norm eigenvectors $\mathbf{u}_2$ corresponding to the eigenvalue $\lambda_2(x, y)$.

Answer

$$\mathbf{u}_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\}$$
$$\mathbf{u}_2 = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$$

Problem 1.5

Give *all* the stationary points of f. For each of them, state whether it is a minimum, a maximum, or a saddle point, and give unit vectors along the directions of maximum and minimum curvature (or state that curvatures in all directions are equal, if that is the case).

Answer

The stationary points are:
maximum point: $(0, 0)$
minimum point: $(2, 2)$
saddle points: $(0, 2)$ and $(2, 0)$
Since the directional curvature of a point is the secondary directional derivative at that point. Let the direction be $\begin{bmatrix} a \\ b \end{bmatrix}$. Therefore:

$$D_u^2 f(x, y) = (a, b) \cdot \left(\frac{\partial}{\partial x}(D_u f(x, y)), \frac{\partial}{\partial y}(D_u f(x, y)) \right)$$
$$D_u^2 f(x, y) = a^2 \frac{\partial^2 f(x, y)}{\partial x^2} + 2ab \frac{\partial^2 f(x, y)}{\partial y \partial x} + b^2 \frac{\partial^2 f(x, y)}{\partial y^2}$$
$$D_u^2 f(x, y) = a^2 (6x - 6) + b^2 (6y - 6)$$

Referring to **Problem 1.6**, $\mathbf{u}_1$ and $\mathbf{u}_2$ are two sets of eigenvectors, which are

$$\mathbf{u}_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\}$$
$$\mathbf{u}_2 = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right\}$$

- For point $(0, 0)$

  $$D_u^2 f(0, 0) = -6a^2 + -6b^2$$

  For both $\mathbf{u}_1$ and $\mathbf{u}_2$,

  $$D_u^2 f(0, 0) = -6$$

  Therefore, the curvatures of point $(0, 0)$ in all directions are equal, which is -6

- Similarly, for point $(2, 2)$

  $$D_u^2 f(2, 2) = 6a^2 + 6b^2$$

For both $\mathbf{u}_1$ and $\mathbf{u}_2$,

$$D_u^2 f(0,0) = 6$$

Therefore, the curvatures of point $(2,2)$ in all directions are equal, which is 6

- For point $(2,0)$

$$D_u^2 f(2,0) = 6a^2 + -6b^2$$

For eigenvectors $\mathbf{u}_1$,

$$c_1 = D_u^2 f(2,0) = 6$$

For eigenvectors $\mathbf{u}_2$,

$$c_2 = D_u^2 f(2,0) = -6$$

Since $c_1 > c_2$, $\mathbf{u}_1$ are the unit vectors along the directions of maximum curvature, and $\mathbf{u}_2$ are the unit vectors along the directions of minimum curvature

- For point $(0,2)$

$$D_u^2 f(0,2) = -6a^2 + 6b^2$$

For eigenvectors $\mathbf{u}_1$,

$$c_1 = D_u^2 f(2,0) = -6$$

For eigenvectors $\mathbf{u}_2$,

$$c_2 = D_u^2 f(2,0) = 6$$

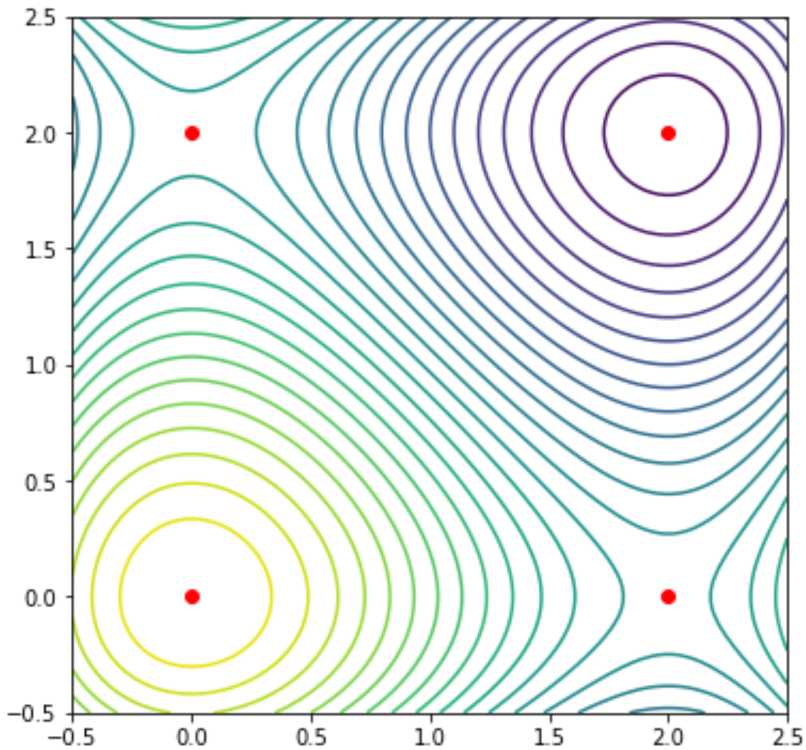
Since $c_2 > c_1$, $\mathbf{u}_2$ are the unit vectors along the directions of maximum curvature, and $\mathbf{u}_2$ are the unit vectors along the directions of minimum curvature

Problem 1.6

To check that your answers are plausible, write Python 3 code that displays a plot of 30 iso-value contours of the function f for $-0.5 \le x \le 2.5$ and $-0.5 \le y \le 2.5$. Draw a red dot at every stationary point.

Answer

In [18]:
```
import numpy
import matplotlib
%matplotlib inline
x = numpy.linspace(-0.5,2.5, num=101)
y = x[:]
xv,yv = numpy.meshgrid(x,y)
z = xv**3 -3*xv**2 + yv**3 -3*yv**2
matplotlib.pyplot.figure(figsize=(6,6))
matplotlib.pyplot.contour (xv,yv,z,30)
matplotlib.pyplot.plot([0,0,2,2],[0,2,0,2],'ro')
matplotlib.pyplot.show()
```



Part 2: Linear Algebra

Problem 2.1

Give four matrices R, N, C, L whose *rows* are bases for the row space, null space, column space (a.k.a. range), and left null space of A. Use rows or columns of A for your basis vectors when possible, and vectors with integer entries in any case.

As a partial check, give the products AN^T and LA (fully spelled out with all their entries).

Answer

$$R = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}, \; N = \begin{bmatrix} 1 & -1 & -1 \end{bmatrix}$$
$$C = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \; L = \begin{bmatrix} -2 & 0 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{bmatrix}$$

To check if the answer is correct, compute:

$$AN^T = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 0 & 2 \\ 1 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0}$$
$$LA = \begin{bmatrix} -2 & 0 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 0 & 2 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{0}$$

Problem 2.2

Give the set S of *all* solutions to the system

$$A\mathbf{x} = \mathbf{0}$$

where A is the matrix above and $\mathbf{0}$ is a column vector of four zeros.

Answer

$$S = \left\{ \begin{bmatrix} v \\ -v \\ -v \end{bmatrix} \right\}, \text{ where } v \in R$$

Problem 2.3

Give the set B of *all* vectors $\mathbf{b}$ for which the system

$$A\mathbf{x} = \mathbf{b}$$

admits a solution, if A is the matrix above.

Answer

$$B = \left\{ \begin{bmatrix} b_1 \\ b_2 \\ 2b_1 \\ b_1 + b_2 \end{bmatrix} \right\}, \text{where } b_1, \ b_2 \in R$$

Part 3: Probability

Problem 3.1

What are the mean m_X and variance σ_X^2 of X?

Answer

$m_X = 0 \times 0.3 + 1 \times 0.5 + 2 \times 0.2 = 0.9$
$\sigma_X^2 = (0 - 0.9)^2 \times 0.3 + (1 - 0.9)^2 \times 0.5 + (2 - 0.9)^2 \times 0.2 = 0.49$

Problem 3.2

What is the mean m_Y of Y?

Answer

$m_Y = \mathbb{E}[\mathbb{E}(Y \mid X)] = \mathbb{E}(Y)$

Let $Z = \mathbb{E}(Y \mid X)$, then, $\mathbb{P}(Z) \ = \ \mathbb{P}(X = x),\ x \in 0,\ 1,\ 2$

$m_Y = \mathbb{E}(Y) = \mathbb{E}(Z) = 6 \times 0.3 + 2 \times 0.5 + 5 \times 0.2 = 3.8$

Problem 3.3

What is the mean $m_{Y|X<2} = \mathbb{E}(Y \mid X < 2)$ of Y given that $X < 2$?

Answer

$$
\begin{aligned}
\mathbb{E}(Y \mid X < 2) &= \int_{-\infty}^{\infty} yf(y|X<2)\ \mathrm{d}y \\
&= \frac{\int_{-\infty}^{\infty} yf(Y=y, X<2)\ \mathrm{d}y}{\mathbb{P}(x<2)} \\
&= \frac{1}{0.8}\left[\int_{-\infty}^{\infty} yf(Y=y, X=0)\ \mathrm{d}y + \int_{-\infty}^{\infty} yf(Y=y, X=1)\ \mathrm{d}y\right] \\
&= \frac{1}{0.8} \times (0.3 \times 6 + 0.5 \times 2) \\
&= 3.5
\end{aligned}
$$

Part 4: Python 3

In [3]:
```
import numpy as np

across = ['marco', 'oneup', 'skate', 'elder', 'seeya']
puzzle = np.array([list(word) for word in across])
print(puzzle)
```

```
[['m' 'a' 'r' 'c' 'o']
 ['o' 'n' 'e' 'u' 'p']
 ['s' 'k' 'a' 't' 'e']
 ['e' 'l' 'd' 'e' 'r']
 ['s' 'e' 'e' 'y' 'a']]
```

Problem 4.1

Write a function with header

```
def prettyPrint(p):
```

that pretty-prints any *nonempty* array like `puzzle` according to the given specifications. Show your code and the result of pretty-printing the given array `puzzle` .

Answer

In [20]:
```
def prettyPrint(p):
    assert p.ndim == 2
    for word in p:
        print(''.join(word))
    print(end='\n')


prettyPrint(puzzle)
prettyPrint(puzzle)
```

```
marco
oneup
skate
elder
seeya

marco
oneup
skate
elder
seeya
```

Problem 4.2

Write four calls to `prettyPrint` that slice the given array `puzzle` to result in the following output:

```
mar
one

der
eya

marco

o
p
e
r
a
```

Your code should comply with the given specifications. Show code and output.

Answer

In [21]:
```
prettyPrint(puzzle[:2, :3])
prettyPrint(puzzle[-2:,-3:])
prettyPrint(puzzle[:1,:])
prettyPrint(puzzle[:, -1:])
```

```
mar
one

der
eya

marco

o
p
e
r
a
```

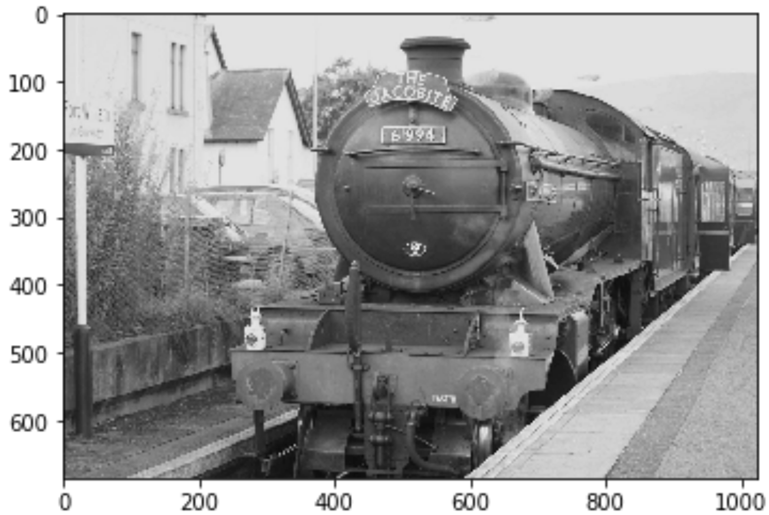
Problem 4.3

In [4]:
```
from skimage import io, color

def readGray(filename):
    img = color.rgb2gray(io.imread(filename))
    return np.around(255 * img).astype(np.uint8)

from matplotlib import pyplot as plt
%matplotlib inline

img = readGray('locomotive.jpg')
plt.imshow(img, cmap='gray')
plt.show()
```



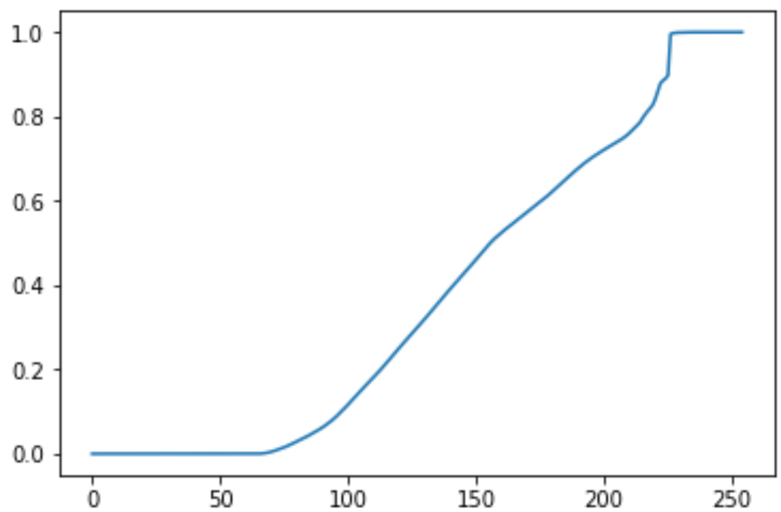
Write a function with header

```
def cumulative(img):
```

that takes an image as produced by `readGray` and computes its cumulative distribution $c(k)$. Show your code, and plot the cumulative histogram of the image in `locomotive.jpg` as read by `readGray` .

Answer

In [5]:
```
import matplotlib.pyplot as plt
def cumulative(img):
    pvalue,pcounts = np.unique(img,return_counts = True)
    pcounts = pcounts/img.size
    pdict = dict(zip(pvalue, pcounts))
    pvalue2 = []
    for i in range(0,256):
        if i not in pvalue:
            pvalue2.append(0)
        if i in pvalue:
            pvalue2.append(pdict[i])
    for i in range(1,256):
        pvalue2[i] = pvalue2[i]+pvalue2[i-1]
    return pvalue2
pvalue2 = cumulative(img)
plt.plot(np.linspace(0,255,256,endpoint = False), pvalue2)
plt.show()
```



Problem 4.4

If your cumulative distribution is correct, the plot should start with several values $c(k)$ equal to zero.

What do these zero values mean? In particular, how do these zeros manifest themselves visually in the image?

Answer

These zero values mean that no pixels in the image have the corresponding intensities. For an example, since $c(0) = 0$, no pixel in the image has 0 intensity. In other words, there is no purely white color in the image. Additionally, since the plot ends with several values of $c(k)$ equal to 1, the intensities except for the first one with $c(k) = 1$ are not used in the image as well.

These zeros and ones imply that the image will have a relatively low contrast comparing to those that utilizes the complete intensity range.

Problem 4.5

Write a function with header

```
def remap(img):
```

that takes an image as returned by `readGray` and replaces the value k of every pixel with value $k' = \text{round}(255\, c(k))$, where the function $c(k)$ was defined in problem 4.3.

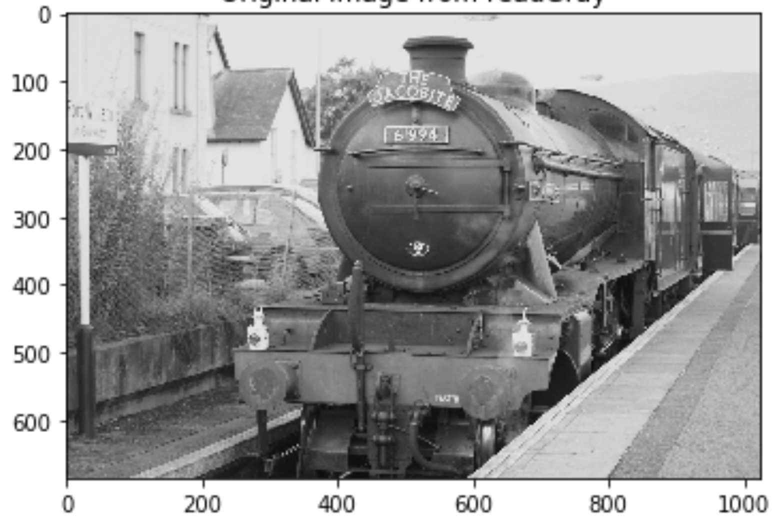
Show your code, display the result of running `remap` on the image in `locomotive.jpg` , and plot the cumulative histogram of the remapped image.

Answer

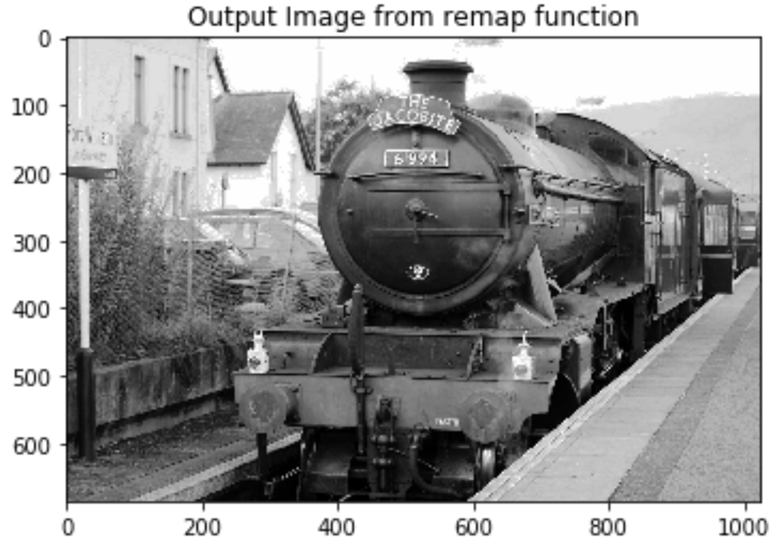
In [24]:

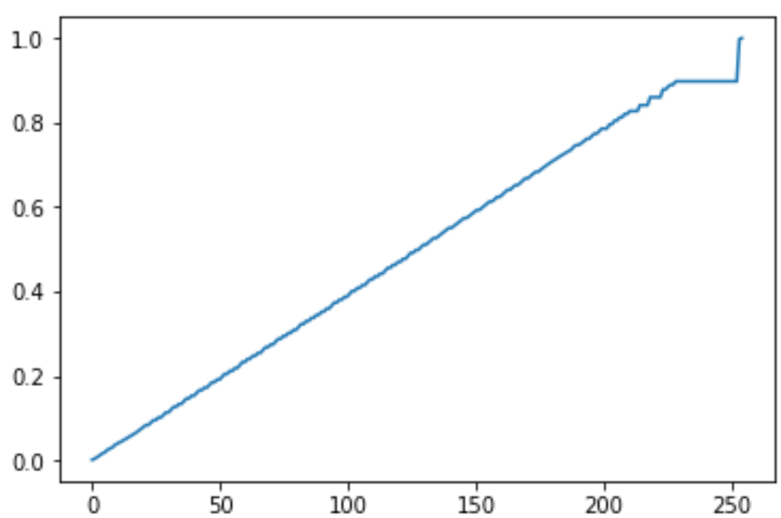
```
row,col = img.shape
img2 = img.flatten()
for i in range(0,row*col):
    img2[i] = round(pvalue2[img2[i]]*255)
img2 = np.reshape(img2,(row,col))
plt.imshow(img.astype(np.uint8),cmap = 'gray')
plt.title('Original Image from readGray')
plt.show()
plt.imshow(img2.astype(np.uint8),cmap = 'gray')
plt.title('Output Image from remap function')
plt.show()
pvalue3 = cumulative(img2)
plt.plot(np.linspace(0,255,256,endpoint = False), pvalue3)
plt.show()
```

Original Image from readGray



Output Image from remap function





Problem 4.6

Comment on the result. In particular, what would the *histogram* (rather than the cumulative distribution) of the remapped image look like, approximately? And in what way does the remapped *image* differ from the original, visually?

Answer

Most of the histogram of the remapped image would be a horizontal line. The intensities between around 225 to 250 would have 0 value in the histogram since the cumulative histogram at those intensities is a straight line. The intensities at the end of the histogram with intensities around 255 will have higher values than the horizontal line because the slope at the end of the cumulative histogram is much larger than the other parts.

Function remap works as contrast stretching. In short, the remapped image will have larger contrast comparing to the original image. Visually, the dark region in the original image will look darker in the remapped image, and the bright region in the original image will look brighter in the remapped image.