Intro ML & AI

Richard D. Desatnik 09/17/2021

Problem 1 Find regression line that fits $(-2,2), (2,4), (3,8), (5,11), (4,7)$

part a

$$\theta_0 = \frac{\sum_{i=1}^{n} (y^i - \theta x^i)}{m} \qquad \theta_1 = \frac{\sum_{i=1}^{n} x^i (y^i - \theta_0)}{\sum_{i=1}^{n} (x^i)^2}$$

$m = 5$

$$\sum_{i=1}^{n} (x^i)^2 = (-2)^2 + 2^2 + 3^2 + 5^2 + 4^2 = 58$$

$$\theta_0 = \frac{(2-\theta_1(-2)) + (4-\theta_1(2)) + (8-\theta_1(3)) + (11-\theta_1(5)) + (17-\theta_1(4))}{58}$$

$$\theta_0 = \frac{2 + \cancel{2\theta} + 4 - \cancel{2\theta} + 8 - 3\theta + 11 - 5\theta + 17 - 4\theta}{5}$$

$$\theta_0 = \frac{42 - 12\theta_1}{5}$$

$$\theta_1 = \frac{-2(2-\theta_0) + 2(4-\theta_0) + 3(8-\theta_0) + 5(11-\theta_0) + 4(17-\theta_0)}{58}$$

$$\theta_1 = \frac{-4 + 2\theta_0 + 8 - 2\theta_0 + 24 - 3\theta_0 + 55 - 5\theta_0 + 68 - 4\theta_0}{58}$$

$$\theta_1 = \frac{151 - 12\theta_0}{58}$$

$$\theta_1 = \frac{151}{58} - \frac{12}{58}(\theta_0)$$

$$\theta_0 = \frac{42}{5} - \frac{12}{5}(\theta_1)$$

$$\theta_1 = \overbrace{\frac{151}{58}}^{r} - \overbrace{\frac{12}{58}}^{s}(\theta_0)$$

$$\theta_0 = \overbrace{\frac{42}{5}}^{p} - \overbrace{\frac{12}{5}}^{q}(\theta_1)$$

$$\theta_1 = \frac{151}{58} - \frac{12}{58}\left(\frac{42}{5} - \frac{12}{5}(\theta_1)\right)$$

$$-\frac{72}{145} + \theta_1 = \frac{151}{58} - \frac{252}{145} + \frac{72}{145}(\theta_1) - \frac{72}{145}(\theta_1)$$

$$\cancel{\frac{145}{73}}\,\theta_1\,\cancel{\frac{73}{145}} = \underbrace{\left(\frac{151}{58} - \frac{252}{145}\right)}_{.8655}\frac{145}{73}$$

$$\theta_1 = 1.719$$

$$\theta_0 = \frac{42}{5} - \frac{12}{5}(1.719)$$

$$\theta_0 = 4.2744$$

$$y = b_0 + b_1 x$$

$$y = 4.2744 + 1.719x$$

$$b_0 = 4.2744$$

$$b_1 = 1.719$$

$$y = 4.27 + 1.719x$$

Problem 1

Part c

$b_0 = 4.0807$

$b_1 = -0.4424$

$$y = 4.0807 - 0.4424x$$

Problem 2
see Matlab at the end

Problem 3

part a

$f(x) = \theta \sigma^{\theta} x^{-\theta-1} \qquad x \geq \sigma$

MLE for σ & θ Expressed in terms of x_i

$x_{min} \geq \sigma$ The MLE of σ is limited to the smallest value of x

lowest value of x

the max possible value of σ is x_{min}

$p(x_i|\theta) = p(x_i|\sigma,\theta)$

$L(\theta) = \prod_{i=1}^{n} p(x_i|\sigma,\theta) = \prod_{i=1}^{n} (\theta\sigma^{\theta}x^{-\theta-1})$

$l(\theta) = \log(L(\theta)) = \log\left(\prod_{i=1}^{n} \theta\sigma^{\theta}x^{-\theta-1}\right)$

$l(\theta) = \sum_{i=1}^{n} \log\left(\theta\sigma^{\theta}x^{-\theta-1}\right)$

$l(\theta) = \sum_{i=1}^{n}\log(\theta) + \sum_{i=1}^{n}\theta\log(\sigma) + \sum_{i=1}^{n}(-\theta\log(x)) - \sum_{i=1}^{n}\log(x)$

$\frac{\partial l}{\partial \theta} = \sum_{i=1}^{n}\frac{1}{\theta} + \sum_{i=1}^{n}\log(\sigma) - \sum_{i=1}^{n}\log(x)$

$0 = \frac{\partial l}{\partial \theta} \qquad 0 = \sum_{i=1}^{n}\frac{1}{\theta} + \sum_{i=1}^{n}\log(\sigma) - \sum_{i=1}^{n}\log(x)$

$\left[\sum_{i=1}^{n}\log(x) - \sum_{i=1}^{n}\log(\sigma)\right]^{-1} = \left(\frac{n}{\theta}\right)^{-1}$

$\cancel{n}\frac{\theta}{\cancel{n}} = \left[\sum_{i=1}^{n}\log(x) - \sum_{i=1}^{n}\log(\sigma)\right]^{-1} n$

$\theta = \frac{n}{\sum_{i=1}^{n}\log(x) - \sum_{i=1}^{n}\log(\sigma)}$

part a continued

$$\Theta = \frac{n}{\sum_{i=1}^{n} \log(x) - \sum_{i=1}^{n} \log(\theta)}$$

$x_i \geq \theta$

$$\Theta = \frac{n}{\sum_{i=1}^{n} \log(x) - \sum_{i=1}^{n} \log(x_i)}$$

x_{min}

$$\Theta = \frac{n}{\sum_{i=1}^{n} \log(x) - n\log(x_i)}$$

MLE for Θ in terms of x_i

ML & AI Hmk 7

Problem 3

b. $x_i = 13, 2, 16, 5, 11, 16, 18, 5, 8, 15$

$\ln(x) \sim N(\mu, \sigma^2) \quad \mu \to \lambda \quad \sigma^2 \to \xi^2$

MLE $\quad p(x|\lambda, \xi^2) = \frac{1}{x \xi \sqrt{2\pi}} \exp\left(\frac{-(\ln(x)-\lambda)^2}{2\xi^2}\right)$

$$L(\theta) = \prod_{i=1}^{n}\left[\frac{1}{x\xi\sqrt{2\pi}} \exp\left(\frac{-(\ln(x)-\lambda)^2}{2\xi^2}\right)\right]$$

$$\ell(\theta) = \log(L(\theta)) = \log\left[\prod_{i=1}^{n} \frac{1}{x\xi\sqrt{2\pi}} \exp\left(\frac{-(\ln(x)-\lambda)^2}{2\xi^2}\right)\right]$$

$$\ell(\theta) = \sum_{i=1}^{n} \log\left(\frac{1}{x\xi\sqrt{2\pi}} \exp\left(\frac{-(\ln(x)-\lambda)^2}{2\xi^2}\right)\right)$$

$$\ell(\theta) = \sum_{i=1}^{n} \log\left(\frac{1}{x\xi\sqrt{2\pi}}\right) + \sum_{i=1}^{n} \log\left[\exp\left(\frac{-(\ln(x)-\lambda)^2}{2\xi^2}\right)\right]$$

$$\ell(\theta) = \underbrace{\cancel{\sum_{i=1}^{n} \log(1)}}_{0} - \sum_{i=1}^{n} \log(x\xi\sqrt{2\pi}) + \sum_{i=1}^{n}\left(\frac{-(\ln(x)-\lambda)^2}{2\xi^2}\right)$$

$$\ell(\theta) = \sum_{i=1}^{n}\left[-\log(x\xi\sqrt{2\pi}) + \left(\frac{-(\ln(x)-\lambda)^2}{2\xi^2}\right)\right]$$

$$\lambda: \quad \frac{\partial\theta}{\partial\lambda} = \frac{\partial\theta}{\partial\lambda}\sum_{i=1}^{n}\left[\underbrace{\cancel{-\log(x\xi\sqrt{2\pi})}}_{0} + \left(\frac{-(\ln(x)-\lambda)^2}{2\xi^2}\right)\right]$$

$$\lambda: \quad \frac{\partial\theta}{\partial\lambda} = \frac{\partial\theta}{\partial\lambda}\sum_{i=1}^{n}\left[\frac{-(\ln(x)-\lambda)^2}{2\xi^2}\right]$$

Problem 3 Part 2

6.
$$\frac{\partial \Theta}{\partial l} = \frac{\partial \Theta}{\partial l} \sum_{i=1}^{n} \left[\frac{-(\ln(x)-\lambda)^2}{2\xi^2} \right]$$

$$0 = \frac{\partial \Theta}{\partial l} \sum_{i=1}^{n} \left[\frac{-(\ln(x)-\lambda)^2}{2\xi^2} \right]$$

$$0 = \frac{-2(\ln(x))}{2\xi^2} \sum_{i=1}^{n} (\ln(x) - \lambda)$$

$$0 = \sum_{i=1}^{n} \ln(x) - \sum_{i=1}^{n} \lambda$$

$$\frac{n\lambda}{n} = \frac{\sum_{i=1}^{n} \ln(x)}{n}$$

$$\lambda = \frac{\sum_{i=1}^{n} \ln(x)}{n}$$

$$\lambda = \frac{\ln(5)+\ln(5)+\ln(16)+\ln(16)+\ln(13)+\ln(18)+\ln(2)+\ln(11)+\ln(8)+\ln(15)}{10}$$

$$\lambda = \frac{14.2194 + 7.8785}{10} = 2.2098$$

b. $\xi : \dfrac{\partial\theta}{\partial\xi} \sum_{i=1}^{n} \left[-\log(x\xi\sqrt{2\pi}) + \left(\dfrac{-(\ln(x)-\lambda)^2}{2\xi^2} \right) \right]$

$$\frac{\partial\theta}{\partial\xi} = \frac{\partial\theta}{\partial\xi} \sum_{i=1}^{n} -\log(x\xi\sqrt{2\pi}) + \frac{\partial\theta}{\partial\xi} \sum_{i=1}^{n} \left(\frac{-(\ln(x)-\lambda)^2}{2\xi^2} \right)$$

$$\frac{\partial\theta}{\partial\xi} = \sum_{i=1}^{n} -\frac{1}{\cancel{x}\,\xi\cancel{\sqrt{2\pi}}}(\cancel{x}\cancel{\sqrt{2\pi}}) + \sum_{i=1}^{n} \left[\frac{\cancel{-}(\ln(x)-\lambda)^2(\cancel{-}\cancel{2})}{\cancel{2}} \right] \xi^{-3}$$

$$\frac{\partial\theta}{\partial\xi} = -\sum_{i=1}^{n} \frac{1}{\xi} + \sum_{i=1}^{n} (\ln(x)-\lambda)^2 \xi^{-3}$$

$$0 = -\sum_{i=1}^{n} \frac{1}{\xi} + \sum_{i=1}^{n} (\ln(x)-\lambda)^2 \xi^{-3}$$

$$\xi^{3} \sum_{i=1}^{n} \frac{1}{\xi} = \left[\sum_{i=1}^{n} (\ln(x)-\lambda)^2 \xi^{-3} \right] \xi^{3}$$

$$\frac{n\xi^2}{n} = \frac{\sum_{i=1}^{n} (\ln(x)-\lambda)^2}{n}$$

$$\xi^2 = \frac{\sum_{i=1}^{n} (\ln(x)-\lambda)^2}{n}$$

1.35433 $\left[(\ln(16) - 2.2098)^2 \times 2 \right.$
$\left. (\ln(5) - 2.2098)^2 \times 2 \right]$

2.46175 $\left[(\ln(13) - 2.2098)^2 \right.$
$(\ln(2) - 2.2098)^2$
$\left. (\ln(11) - 2.2098)^2 \right]$

.7284245 $\left[(\ln(18) - 2.2098)^2 \right.$
$(\ln(8) - 2.2098)^2$
$+ \left. (\ln(15) - 2.2098)^2 \right]$

4.5445

$$\xi^2 = \frac{4.5445}{10}$$

$$\xi^2 = 0.45445$$

Problem 3

part c

$$f(x) = (x;\theta) = \frac{x}{\theta^2} e^{-\frac{x}{\theta}} \qquad x > 0$$

MLE for θ $\qquad \theta > 0$

$$p(x_i|\theta) = \frac{x}{\theta^2} e^{-\frac{x}{\theta}}$$

$$L(\theta) = \prod_{i=1}^{n} \left(\frac{x}{\theta^2} \exp\left(-\frac{x}{\theta}\right)\right)$$

$$l(\theta) = \log(L(\theta)) = \log\left(\prod_{i=1}^{n} \frac{x}{\theta^2} \exp\left(-\frac{x}{\theta}\right)\right)$$

$$l(\theta) = \sum_{i=1}^{n} \log\left(\frac{x}{\theta^2}\right) + \sum_{i=1}^{n} \log\left(\exp\left(-\frac{x}{\theta}\right)\right)$$

$$l(\theta) = \sum_{i=1}^{n} \log(x) - \sum_{i=1}^{n} \log(\theta^2) + \sum_{i=1}^{n} -\frac{x}{\theta}$$

$$l(\theta) = \sum_{i=1}^{n} \log(x) - \sum_{i=1}^{n} 2\log(\theta) + \sum_{i=1}^{n} -x\theta^{-1}$$

$$\frac{\partial\theta}{\partial l} = -\sum_{i=1}^{n} \frac{2}{\theta} + \sum_{i=1}^{n} x\theta^{-2}$$

$$0 = -\sum_{i=1}^{n} \frac{2}{\theta} + \sum_{i=1}^{n} x\theta^{-2}$$

$$\sum_{i=1}^{n} \frac{2}{\theta} = \sum_{i=1}^{n} x\theta^{-2}$$

$$\theta^{\cancel{2}} \frac{2n}{\cancel{\theta}} = \sum_{i=1}^{n} x\theta^{-2} \; \theta^2$$

$$\frac{2n\theta}{2n} = \frac{\sum_{i=1}^{n} x}{2n}$$

MLE : θ $\qquad \theta = \dfrac{\sum_{i=1}^{n} x}{2n}$

MLAI_Homework1_FinalCode

September 18, 2021

[1]:
```
#Problem1_PartB

import csv
i = 0
xlist= []
ylist= []
sumxylist = []
xsquarelist = []
ysquarelist = []
with open('C:\\Users\\rdesa\\OneDrive\\Desktop\\MLAI_Hmk1\\MLAI_Hmkb2.csv') as␣
 ↪csv_file:
    csv_reader = csv.reader(csv_file, delimiter=',')
    for row in csv_reader:
        if row[0] == 'X':
            print("skip")
        else:
            x = float(row[0])
            y = float(row[1])
            xy = x*y
            xsquare = x**2
            ysquare = y**2
            xlist.append(float(row[0]))
            ylist.append(float(row[1]))
            sumxylist.append(x*y)
            xsquarelist.append(x**2)
            ysquarelist.append(y**2)

n = len(xlist)
xsum = sum(xlist)
ysum = sum(ylist)
xylistsum = sum(sumxylist)
xsquaresum = sum(xsquarelist)
p = (ysum/n)
q = (xsum/n)
r = (xylistsum)/xsquaresum
s = xsum/xsquaresum
b1 = (r-(s*p))/(1-(s*q))
```

```
b0 = p-(q*b1)
print("This is b1: " + str(b1))
print("This is b0: " + str(b0))
print("Solution: y=" + str(b0) + "+" + str(b1) + "x")
```

```
skip
This is b1: 1.7191780821917806
This is b0: 4.273972602739727
Solution: y=4.273972602739727+1.7191780821917806x
```

[2]:

```
#Problem1_PartC
import csv
i = 0
xlist= []
ylist= []
sumxylist = []
xsquarelist = []
ysquarelist = []
#with open('C:\\Users\\rdesa\\OneDrive\\Desktop\\Recitation 1\\data.csv') as␣
↪csv_file:
with open('C:\\Users\\rdesa\\OneDrive\\Desktop\\Assigment_1 (2)\\data.csv') as␣
↪csv_file:
    csv_reader = csv.reader(csv_file, delimiter=',')
    for row in csv_reader:
        if row[0] == 'X':
            print("skip")
        else:
            x = float(row[0])
            y = float(row[1])
            xy = x*y
            xsquare = x**2
            ysquare = y**2
            xlist.append(float(row[0]))
            ylist.append(float(row[1]))
            sumxylist.append(x*y)
            xsquarelist.append(x**2)
            ysquarelist.append(y**2)

n = len(xlist)
xsum = sum(xlist)
ysum = sum(ylist)
xylistsum = sum(sumxylist)
xsquaresum = sum(xsquarelist)
ysquaresum = sum(ysquarelist)
p = (ysum/n)
q = (xsum/n)
r = (xylistsum)/xsquaresum
```

```
s = xsum/xsquaresum
b1 = (r-(s*p))/(1-(s*q))
b0 = p-(q*b1)
print("This is b1: " + str(b1))
print("This is b0: " + str(b0))
print("Solution: y=" + str(b0) + "+" + str(b1) + "x")
```

```
skip
This is b1: -0.44236913850438125
This is b0: 4.080657141896113
Solution: y=4.080657141896113+-0.44236913850438125x
```

[3]:
```
#Problem2_Parta
def includes (item, val, start_ind=0):
    if type(item) == list:
        #print("list")
        if val in item[start_ind:]:
            print(True)
        else:
            print(False)
    elif type(item) == str:
        #print("string")
        if val in item[start_ind:]:
            print(True)
        else:
            print(False)
    elif type(item) == dict:
        #print("dictionary")
        if val in item.values():
            print(True)
        else:
            print(False)
    else:
        print("Please change your item, as it is not a string, list, or␣
 ↪dictionary")

#item = [1,2,3]
#val = 1
includes([2,3,4],2)
includes([2,3,4],4)
includes({'a':1,'b':2},1)
includes({'a':1,'b':2},'a')
includes('abcd','b')
```

```
True
True
True
```

```
False
True
```

[4]:
```
#Problem2_Partb
avglist = []
def moving_average(newvalue):
    avglist.append(newvalue)
    n = len(avglist)
    total = sum(avglist)
    avg = (total/n)
    navg = round(avg,1)
    return navg

mAvg = print(moving_average(10))
#print(avglist)
mAvg = print(moving_average(11))
#print(avglist)
mAvg = print(moving_average(12))
#print(avglist)
```

```
10.0
10.5
11.0
```

[5]:
```
#Problem2_Partc
def same_frequency(num1,num2):
    if type(num1) != int or type(num2) != int:
        print("please use numbers only")
    else:
        print("Cool, you put in two numbers")
        num1 = str(num1)
        num2 = str(num2)
        dictnum1 = {"1": num1.count("1"), "2": num1.count("2"), "3":num1.
 ↪count("3"), "4":num1.count("4"), "5":num1.count("5"), "6":num1.count("6"), "7":
 ↪num1.count("7"), "8":num1.count("8"), "9":num1.count("9"), "0":num1.count("0")}
        dictnum2 = {"1": num1.count("1"), "2": num2.count("2"), "3":num2.
 ↪count("3"), "4":num2.count("4"), "5":num2.count("5"), "6":num2.count("6"), "7":
 ↪num2.count("7"), "8":num2.count("8"), "9":num2.count("9"), "0":num2.count("0")}
        if dictnum1 == dictnum2:
            print(True)
        else:
            print(False)
same_frequency(552211,221515)
same_frequency(321142,3212215)
same_frequency(12345,31354)
same_frequency(1212,2211)
same_frequency(2,"two")
```

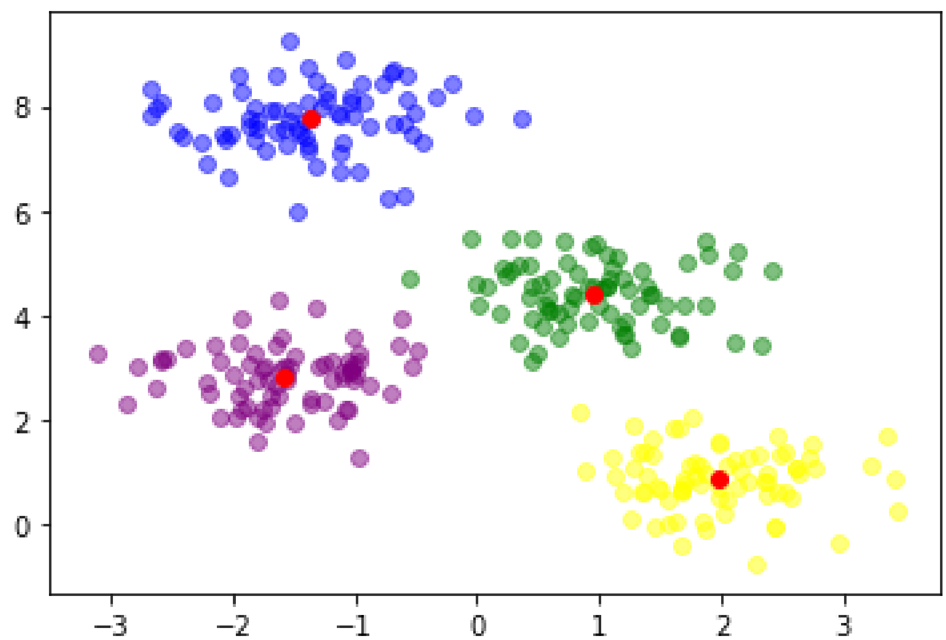
```
Cool, you put in two numbers
True
Cool, you put in two numbers
False
Cool, you put in two numbers
False
Cool, you put in two numbers
True
please use numbers only
```

[9]:

```python
#Problem2_Partd
import numpy as np
import matplotlib.pyplot as plt
i=0
c=0
n=0
m=0
Doc = np.load("C:\\Users\\rdesa\\OneDrive\\Desktop\\Assigment_1 (2)\\kmeans.npz")
DATA = Doc['data']
PRED = Doc['pred']
CENTERS = Doc['centers']
xblist = []
yblist = []
xylist = []
yylist = []
xglist = []
yglist = []
xprlist = []
yprlist = []
xclist = []
yclist = []
xplist = []
for value in PRED:
    xplist.append(value)

for value in DATA[0:]:
    if xplist[i] == 0:
        xblist.append(float(value[0]))
        yblist.append(float(value[1]))
    else:
        None
    i = i+1
for value in DATA[0:]:
    if xplist[n] == 1:
        xylist.append(float(value[0]))
        yylist.append(float(value[1]))
    else:
```

```
        None
    n = n + 1
for value in DATA[0:]:
    if xplist[m] == 2:
        xglist.append(float(value[0]))
        yglist.append(float(value[1]))
    else:
        None
    m = m+1
for value in DATA[0:]:
    if xplist[c] == 3:
        xprlist.append(float(value[0]))
        yprlist.append(float(value[1]))
    else:
        None
    c = c+1
for value in CENTERS[0:]:
    xclist.append(float(value[0]))
    yclist.append(float(value[1]))

#print(len(xlist))
#print(len(ylist))

plt.scatter(xblist,yblist, c = "blue", alpha = 0.5)
plt.scatter(xylist,yylist, c = "yellow", alpha = 0.5)
plt.scatter(xglist,yglist, c = "green", alpha = 0.5)
plt.scatter(xprlist,yprlist, c="purple", alpha = 0.5)
plt.scatter(xclist,yclist, c="red")
```

[9]: <matplotlib.collections.PathCollection at 0x1b45cd69b88>



[7]:
```
#Problem2_Parte
import numpy as np
#v = np.array([1,3,8,7])
#u = np.array([2,9,6,5])
np.random.seed(24787)
X = np.random.randint(-1000,1000, size=3000)
Y = np.random.randint(-1000,1000, size=3000)
def NUMPY_outer(v,u):
    (lenx,) = v.shape
    (leny,) = u.shape
    ONE = np.ones((lenx,leny))
    unew = np.reshape(v,(1,leny))
    utranspose = unew.T
    check = np.multiply(u,ONE)
    finalanswer = np.multiply(check,utranspose)
    print(finalanswer)
    return finalanswer

result = NUMPY_outer(X,Y)
#print(result)
```

```
[[ 288116.  433466.  322354. ...  234498.  459306.  323646.]
 [ 214972.  323422.  240518. ...  174966.  342702.  241482.]
 [-312200. -469700. -349300. ... -254100. -497700. -350700.]
 ...
 [ 180184.  271084.  201596. ...  146652.  287244.  202404.]
```

```
[ -66454.  -99979.  -74351. ...  -54087. -105939.  -74649.]
[ 203376.  305976.  227544. ...  165528.  324216.  228456.]]
```

[8]:
```
#Checking problem 2e
np.random.seed(24787)
X = np.random.randint(-1000,1000, size=3000)
Y = np.random.randint(-1000,1000, size=3000)
np.outer(X,Y)
```

[8]:
```
array([[ 288116,  433466,  322354, ...,  234498,  459306,  323646],
       [ 214972,  323422,  240518, ...,  174966,  342702,  241482],
       [-312200, -469700, -349300, ..., -254100, -497700, -350700],
       ...,
       [ 180184,  271084,  201596, ...,  146652,  287244,  202404],
       [ -66454,  -99979,  -74351, ...,  -54087, -105939,  -74649],
       [ 203376,  305976,  227544, ...,  165528,  324216,  228456]])
```

[ ]:

[ ]: