

In [1]:

```
import numpy as np
import matplotlib.pyplot as plt
```

In [2]:

```
##### CSE 546 #####
##### HW 2 #####
##### MUDI QIN #####
```

In [3]:

```
##### LASSO SET UP #####
```

In [4]:

```

def CoordinateDescent(y,X, Lambda,tolerance,w_init=None):
    '''
    n is the number of record
    d is

    y is a n-by-1 vector

    X is composed of [x1;x2;x3...;xn], where xn is 1-by-d

    w is is a d-by-1 vector

    '''
    n,d = X.shape
    b=0
    ak=0
    ck = 0
    ek = 1000;
    w_curr=np.zeros((d,));
    loop_count =0
    print("+++++++Calculating for Lambda = {}+++++++".format(Lambda))
    while (ek > tolerance):
        if(loop_count > 10000):
            print("didn't converge")
            break
        loop_count += 1
        w_prev = np.copy (w_curr)
        b0=np.dot(w_prev.T,X.T)
        #print("the shape of b0 is {}".format(b0.shape))
        c = np.zeros((d,))
        b = 1/n * (np.sum(y -b0))
        #print("the shape of b is {}".format(b.shape))
        #print("b is {}".format(b))
        #print("b shape {}".format(b.shape))
        a = 2*np.sum(np.square(X), axis=0)
        for k in range(0,d):
            selector = [j for j in range(d) if j != k]
            #p1=np.dot(X[:, selector], w_curr[selector])+b
            #p2 = y-p1
            #c[k] = 2 * np.dot(X[:,k], p2 )
            c[k] = 2*np.dot(X[:, k], y - (b + np.dot(w_curr.T, X.T) - w_curr[k]*X
           [:, k]))
            #print("the shape of ck is {}".format(ck.shape))
            #print(ck)
            #print(ck.shape)
            if (c[k] < -Lambda):
                w_curr[k] = (c[k]+ Lambda) / a[k]
            elif (c[k] > Lambda):
                w_curr[k] = (c[k]- Lambda) / a[k]
            else:
                w_curr[k] = 0
            #if (w_curr[k]!= 0):
            #print(w_curr[k])

        #print(np.linalg.norm(w_curr))

```

```

        #print(np.linalg.norm(w_prev))
        ek = np.max(np.abs(w_curr - w_prev))
        print("..... it takes {} iterations to converg
e.....".format(loop_count))
        print(".....error is {}".format(ek))
        #Lambda = Lambda/2
        return w_curr

```

In [5]:

```

#####
##### A 4 #####
#####

```

In [6]:

```

##### setting up parameters #####
n = 500; d= 1000; k = 100; sigma =1
w1 = np.array([i/k for i in range(1,k+1)])
w2 = np.zeros(900,dtype=w1.dtype);
w_defined=np.r_[w1,w2]

```

In [7]:

```

##### generate record X #####
np.random.seed(321)
X = np.random.normal(size = (n,d))
##### generte y #####
y = np.dot(w_defined.T,X.T) + np.random.normal(scale = np.sqrt(sigma), size=(n,))
print(y.shape)

```

(500,)

In [8]:

```

##### LambdaMax #####33
Lambda = 2*np.max( np.abs(np.dot(X.T, (y-np.mean(y)).T)))

```

In []:

```

#####
##### A4 a #####
#####

```

In [9]:

```
nonzero_entries = 0;
zero_entries = d;
lambda_zeros = {}
W = list()
while (zero_entries != 0):
    w = CoordinateDescent(y,X,Lambda,1E-4)
    W.append(w)
    zero_entries = np.count_nonzero(w==0)
    nonzero_entries = d - zero_entries
    lambda_zeros.update({Lambda:nonzero_entries})
    print("non-zero entries count : {}".format(nonzero_entries))
    Lambda = Lambda/2
```

```
+++++Calculating for Lambda = 1276.04410217830
5+++++
..... it takes 1 iterations to converg
e.....
.....error is 0.0.....
non-zero entries count : 0
+++++Calculating for Lambda = 638.022051089152
5+++++
..... it takes 7 iterations to converg
e.....
.....error is 4.5659093030211984e-05.....
non-zero entries count : 41
+++++Calculating for Lambda = 319.011025544576
2+++++
..... it takes 10 iterations to converg
e.....
.....error is 5.931050070814847e-05.....
non-zero entries count : 80
+++++Calculating for Lambda = 159.505512772288
1+++++
..... it takes 12 iterations to converg
e.....
.....error is 7.914814171627871e-05.....
non-zero entries count : 126
+++++Calculating for Lambda = 79.7527563861440
6+++++
..... it takes 19 iterations to converg
e.....
.....error is 7.482086623546852e-05.....
non-zero entries count : 199
+++++Calculating for Lambda = 39.8763781930720
3+++++
..... it takes 36 iterations to converg
e.....
.....error is 8.503278681352977e-05.....
non-zero entries count : 283
+++++Calculating for Lambda = 19.9381890965360
14+++++
..... it takes 62 iterations to converg
e.....
.....error is 9.281536839100823e-05.....
non-zero entries count : 371
+++++Calculating for Lambda = 9.96909454826800
7+++++
..... it takes 135 iterations to converg
e.....
.....error is 9.884973574977862e-05.....
non-zero entries count : 421
+++++Calculating for Lambda = 4.98454727413400
4+++++
..... it takes 233 iterations to converg
e.....
.....error is 9.827994385169792e-05.....
non-zero entries count : 468
+++++Calculating for Lambda = 2.49227363706700
2+++++
..... it takes 319 iterations to converg
```

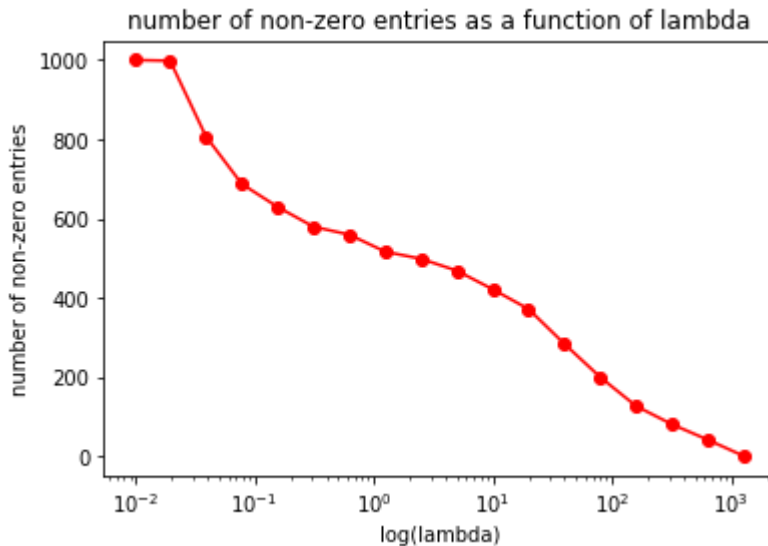
```
e.....
.....error is 9.987048033764656e-05.....
non-zero entries count : 498
+++++Calculating for Lambda = 1.24613681853350
1+++++
..... it takes 508 iterations to converg
e.....
.....error is 9.961901241699322e-05.....
non-zero entries count : 516
+++++Calculating for Lambda = 0.62306840926675
04+++++
..... it takes 696 iterations to converg
e.....
.....error is 9.99463305993431e-05.....
non-zero entries count : 559
+++++Calculating for Lambda = 0.31153420463337
52+++++
..... it takes 1007 iterations to converg
e.....
.....error is 9.986650269471278e-05.....
non-zero entries count : 579
+++++Calculating for Lambda = 0.15576710231668
76+++++
..... it takes 1279 iterations to converg
e.....
.....error is 9.935870902033628e-05.....
non-zero entries count : 629
+++++Calculating for Lambda = 0.07788355115834
38+++++
..... it takes 1594 iterations to converg
e.....
.....error is 9.98916073887135e-05.....
non-zero entries count : 687
+++++Calculating for Lambda = 0.03894177557917
19+++++
..... it takes 1381 iterations to converg
e.....
.....error is 9.983385581144397e-05.....
non-zero entries count : 807
+++++Calculating for Lambda = 0.01947088778958
595+++++
..... it takes 22 iterations to converg
e.....
.....error is 9.948655049119015e-05.....
non-zero entries count : 998
+++++Calculating for Lambda = 0.00973544389479
2976+++++
..... it takes 18 iterations to converg
e.....
.....error is 7.417144001875542e-05.....
non-zero entries count : 1000
```

In [14]:

```

lambdas = list(lambda_zeros.keys())
nonzero_count = list(lambda_zeros.values())
plt.plot(lambdas, nonzero_count, 'r-o')
plt.xscale('log')
plt.yscale('linear')
plt.xlabel("log(lambda)")
plt.ylabel("number of non-zero entries")
plt.title("number of non-zero entries as a function of lambda")
plt.savefig('A4_a_linearY')
plt.show()

```



In [64]:

```
lambda_zeros.keys()
```

Out[64]:

```
dict_keys([1276.044102178305, 638.0220510891525, 319.0110255445762, 15
9.5055127722881])
```

In []:

```

#####
##### A4 b #####
#####

```

In [11]:

```
##### FDR #####
##
nonzero_count = np.array(nonzero_count)
incorrect_nonzero = np.array([np.count_nonzero(x[k:]) for x in W])
FDR = np.divide(incorrect_nonzero, nonzero_count)
FDR
```

/home/mumu/anaconda3/lib/python3.7/site-packages/ipykernel_launcher.p
y:4: RuntimeWarning: invalid value encountered in true_divide
after removing the cwd from sys.path.

Out[11]:

```
array([      nan, 0.14634146, 0.225      , 0.4047619 , 0.54271357,
        0.66784452, 0.74393531, 0.7695962 , 0.78846154, 0.80321285,
        0.81007752, 0.82289803, 0.82901554, 0.84260731, 0.8558952 ,
        0.87608426, 0.8997996 , 0.9       ])
```

In [12]:

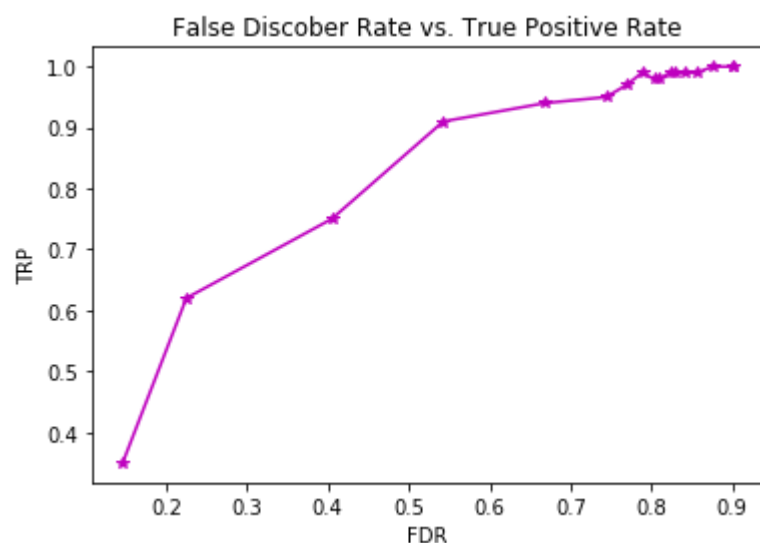
```
##### TPR #####
##
TPR = np.array([np.count_nonzero(x[:k]) for x in W])/k
TPR
```

Out[12]:

```
array([0.   , 0.35, 0.62, 0.75, 0.91, 0.94, 0.95, 0.97, 0.99, 0.98, 0.9
      8,
        0.99, 0.99, 0.99, 0.99, 1.   , 1.   , 1.   ])
```


In [15]:

```
##### plot FDR, TPR #####  
####  
plt.plot(FDR,TPR, 'm-*')  
  
plt.xlabel("FDR")  
plt.ylabel("TRP")  
plt.title("False Discober Rate vs. True Positive Rate")  
plt.savefig('A4_b')  
plt.show()
```



In []: