Jintack Lim
jl4312
Homework 3, Problem 1

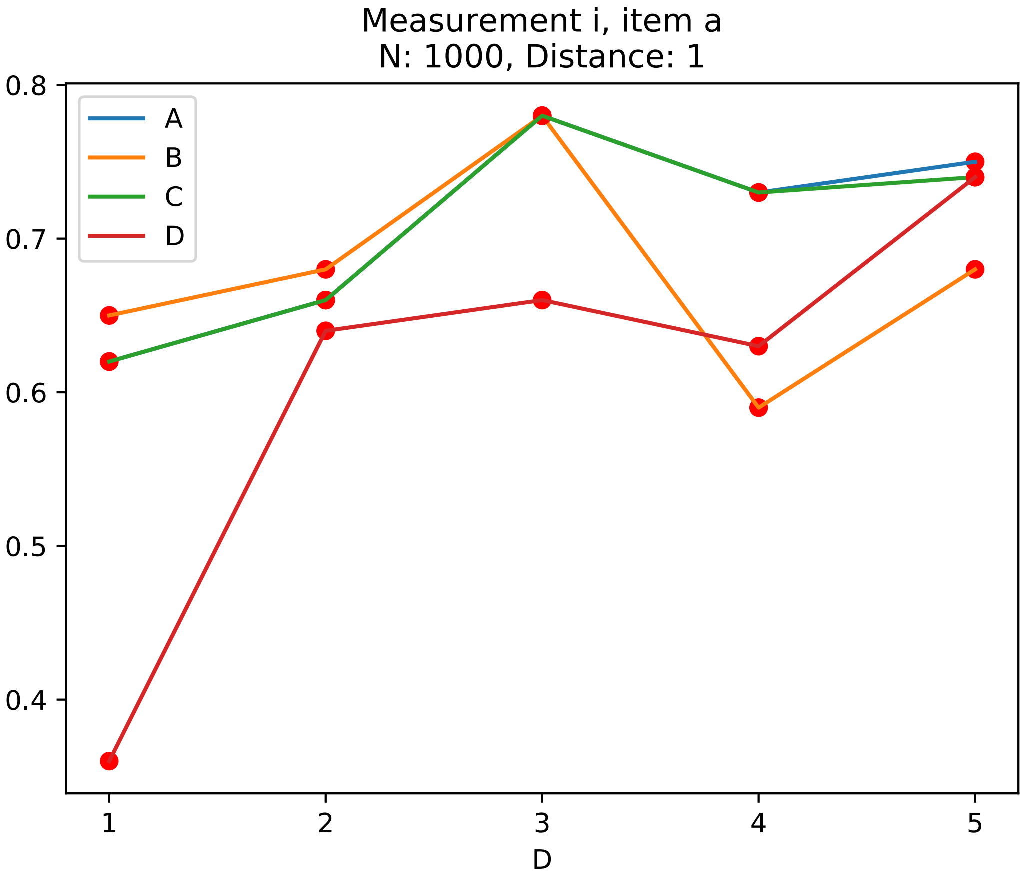
1) -1/+1 or 0/1
Theoretically, we need to use -1/+1 for perceptron and svm, and use 0/1 for logistic regression. However, sklearn LogisticRegression and LinearSVC can handle 0/1 without problem [1]. In my perceptron implementation of classifier D, I used -1/+1.

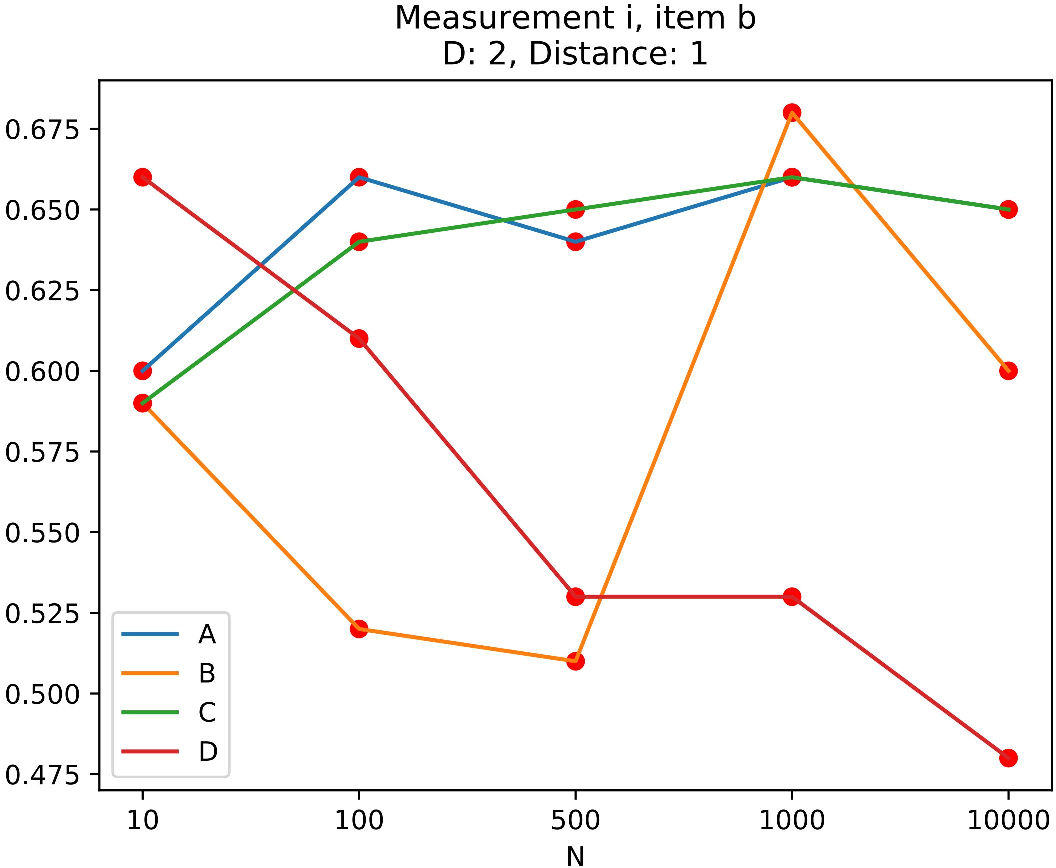
2) Unseparable data
In classifier D, I used fixed number of iterations to avoid infinite execution in case of handling unseparable data.

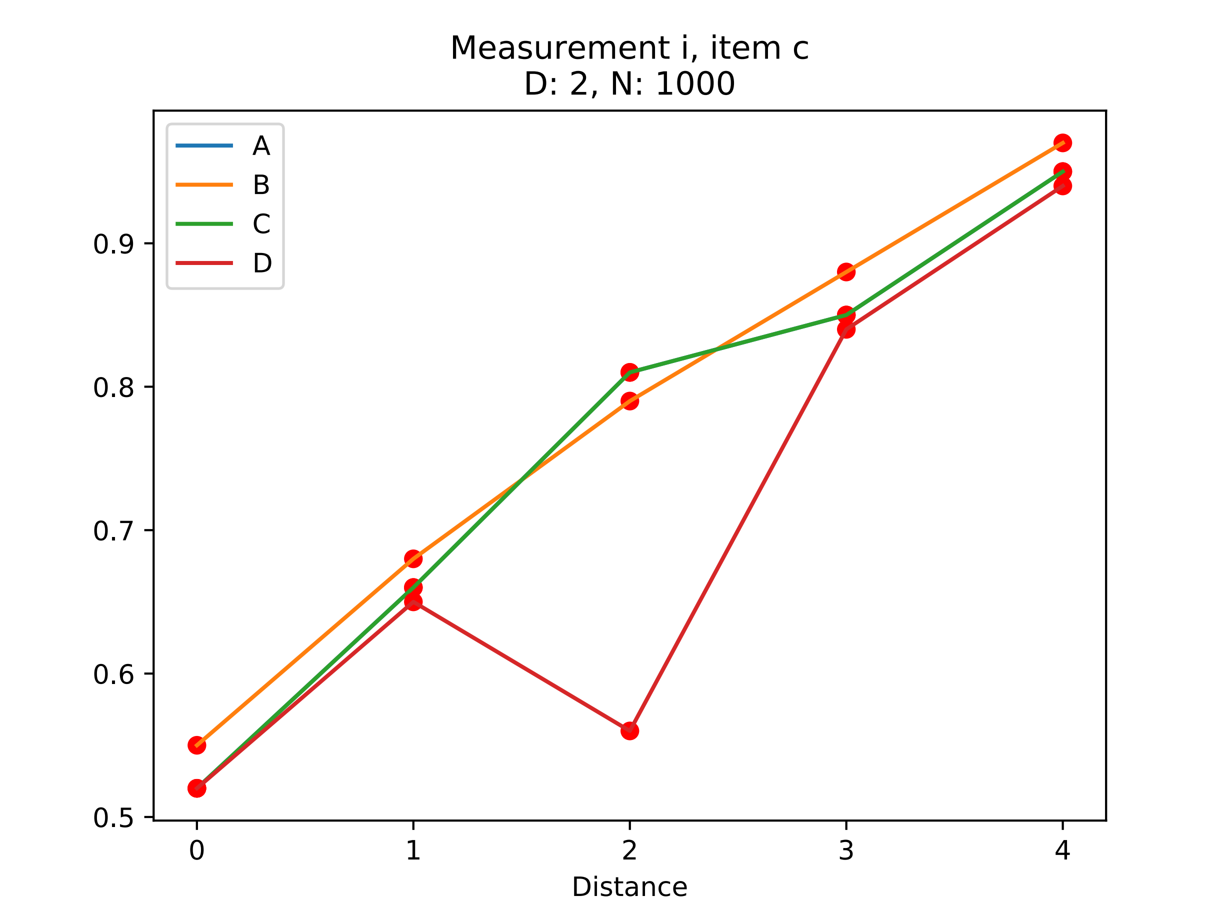
3) N, D, Distance
In this assignment, Distance is the key variable.
If the distance is 0, then prediction using ANY classifier would be around 0.5.
If the distance is more than 3, then prediction would be close to 1.
So, I picked distance 1 for the fixed value for item a and b. For item c, I changed it from 0 to 4.

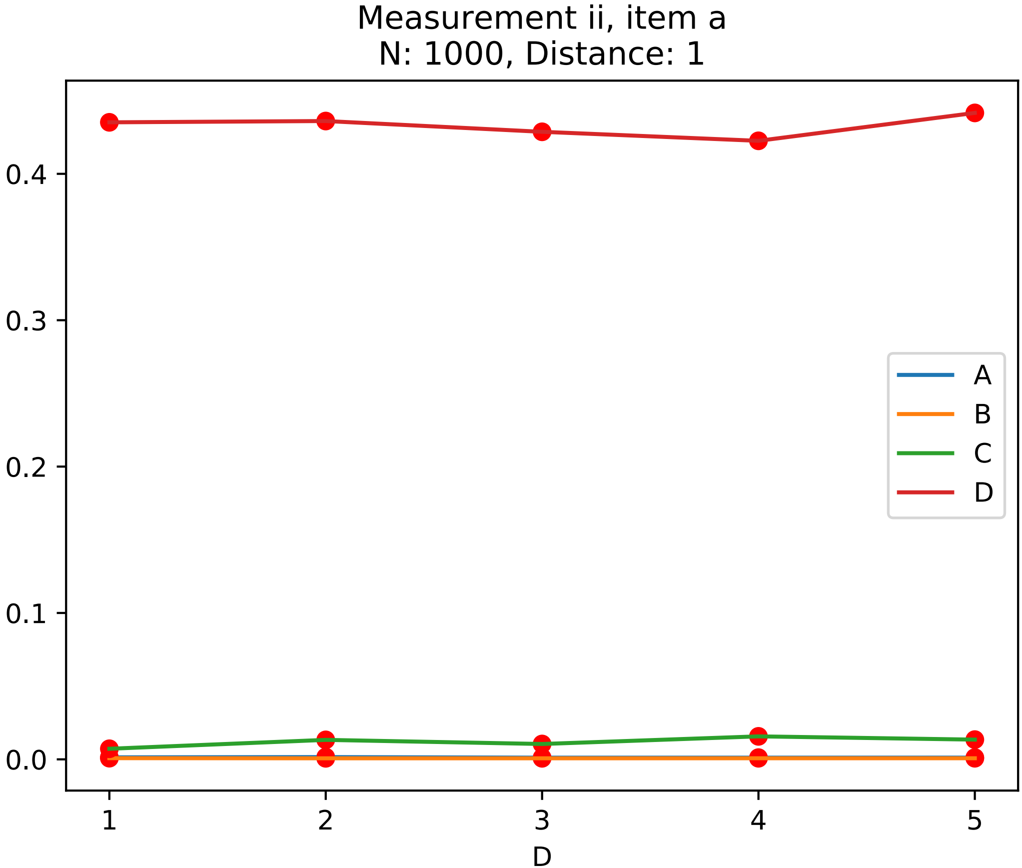
N and D are related to overfit problem.
For normal cases where N is much larger than D, then there's no problem.
However if D gets close to N, then there could be overfit problem.

[1] http://stamfordresearch.com/scikit-learn-perceptron/

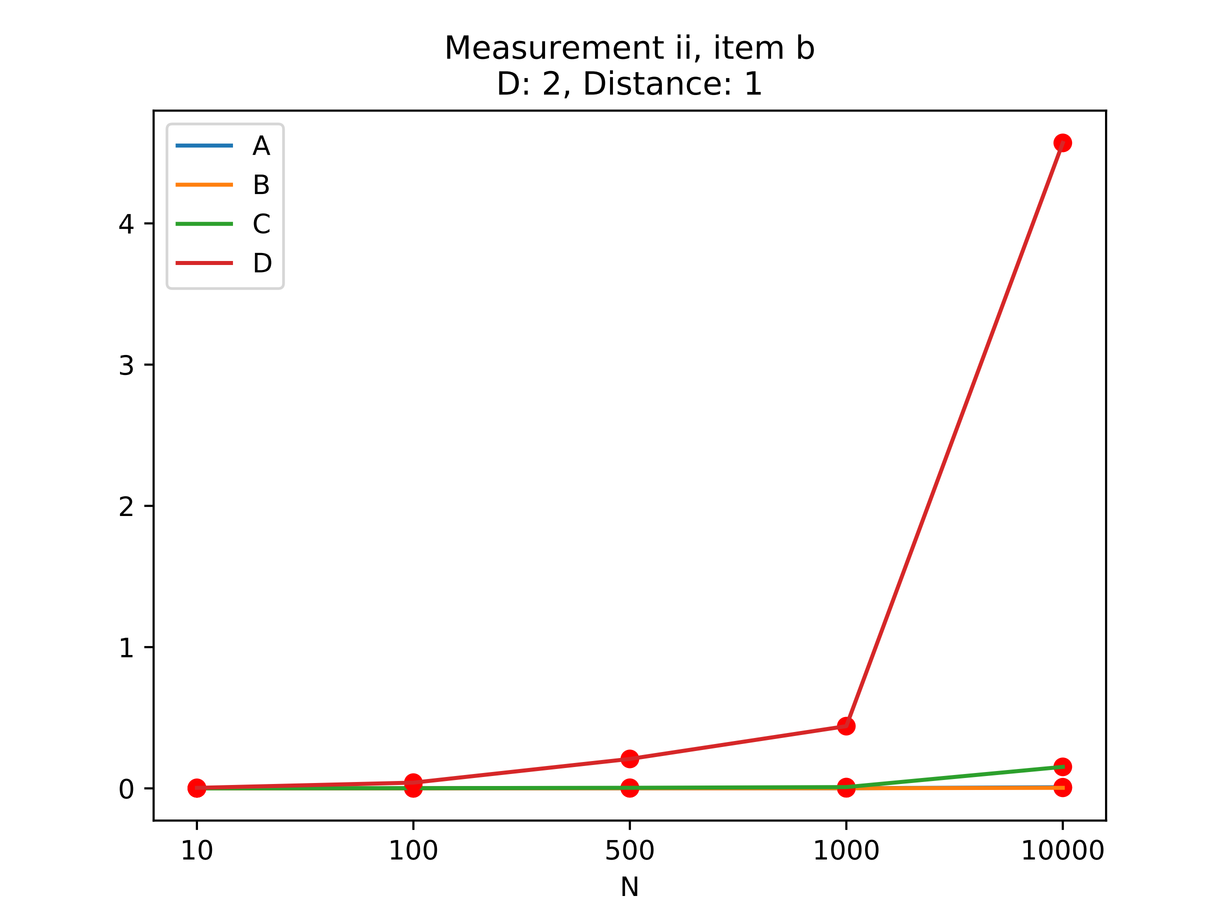

Measurement i, item a
N: 1000, Distance: 1
A
B
C
D
0.8
0.7
0.6
0.5
0.4
1
2
3
4
5
D


Measurement i, item b
D: 2, Distance: 1
0.675
0.650
0.625
0.600
0.575
0.550
0.525
0.500
0.475
10
100
500
1000
10000
N
A
B
C
D

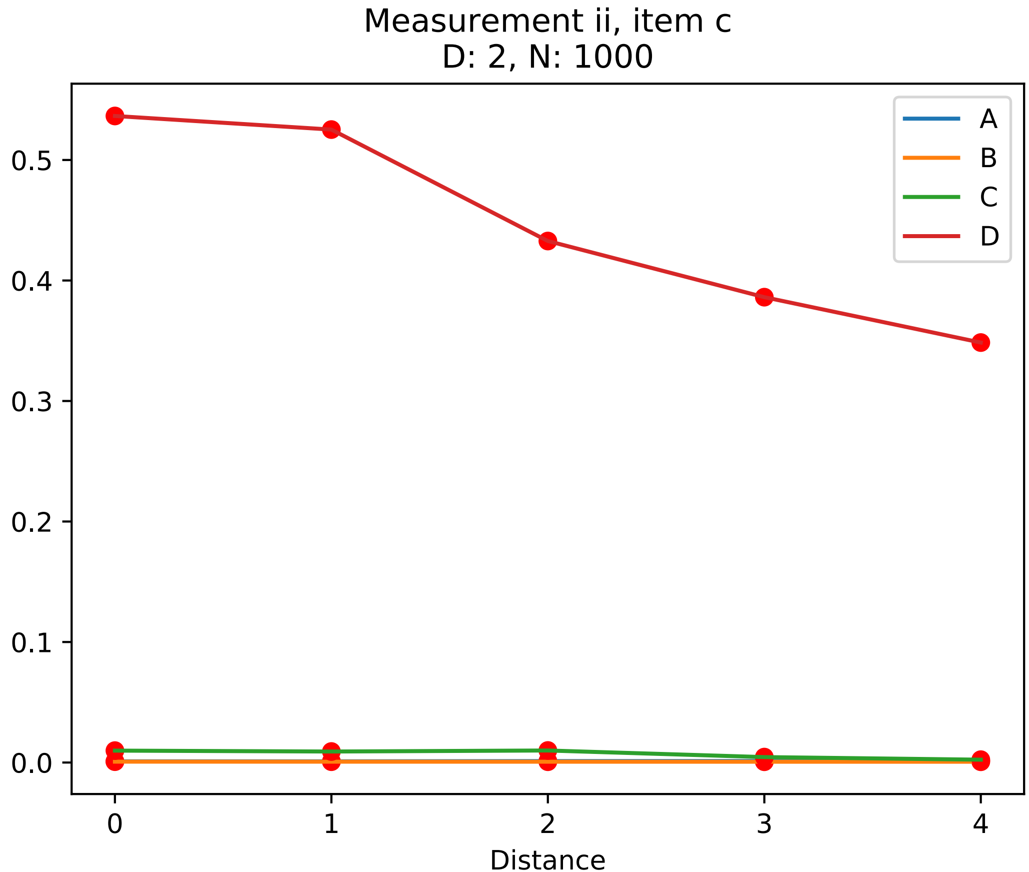

Measurement i, item c
D: 2, N: 1000
A
B
C
D
0.5
0.6
0.7
0.8
0.9
0
1
2
3
4
Distance



Measurement ii, item a
N: 1000, Distance: 1
0.4
0.3
0.2
0.1
0.0
1
2
3
4
5
D
A
B
C
D



Measurement ii, item b
D: 2, Distance: 1
A
B
C
D
0
1
2
3
4
10
100
500
1000
10000
N



Measurement ii, item c
D: 2, N: 1000
0.5
0.4
0.3
0.2
0.1
0.0
0
1
2
3
4
Distance
A
B
C
D

COMS 4771 Spring 2017 Homework 3

Jin Tack Lim, jl4312

Problem 2
(a)
Constraint: $\sum_{i=d}^{D} x_d = C$
If we take derivatives for each dimension d,
$\frac{d}{x_d} = \lambda$
We can plug this to the Constraint equation, and get
$\frac{1}{\lambda} = \frac{2C}{D(D+1)}$
Therefore
$x_d = \frac{2Cd}{D(D+1)}$
and Maximum value of f(x) is

$$\boxed{\sum_{d=1}^{D} dlog\frac{2Cd}{D(D+1)}}$$

(b)
Constraint: $\sum_{d=1}^{D} x_i^2 \leq 1$
So, these are all equations and inequalities we have.
$\frac{1}{x_d} = \lambda 2x_d$ for each dimension d.
$\sum_{d=1}^{D} x_d^2 \leq 1$
$\lambda(\sum_{i=d}^{D} x_d^2 - 1) = 0$
From the first equation, if λ is zero, then all x_i goes to infinity which does not satisfy the second inequality. So $\sum_{d=1}^{D} x_d^2 = 1$.
Then, by solving above equations,
$x_d = \sqrt{\frac{2d}{D(D+1)}}$
and Maximum value of f(x) is

$$\boxed{\sum_{d=1}^{D} dlog\sqrt{\frac{2d}{D(D+1)}}}$$

Howework 3 Problem 3
Jintack Lim
jl4312

a) if H is NOT positive definite, there could be no solution at all. However this is not sufficient condition.
b) if H is NOT positive definite, there could be no finite solution. However this is not sufficient condition.
c) there will be a single solution if H is positive definite and C has full rank.
d) there will be a single solution if H is positive definite and C does not have full rank.