CS 541-A Artificial Intelligence: Final Exam

Instructor: Jie Shen

12/15/2020, 6:30 pm - 9:00 pm EST

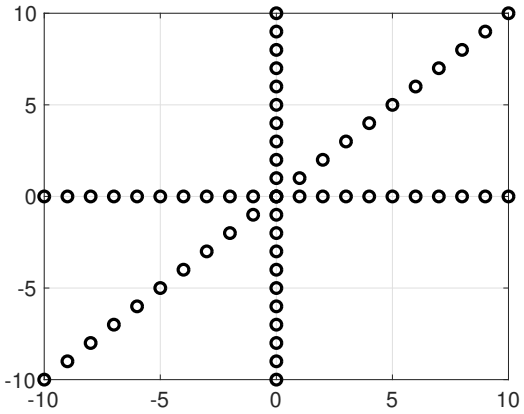
Instructions:

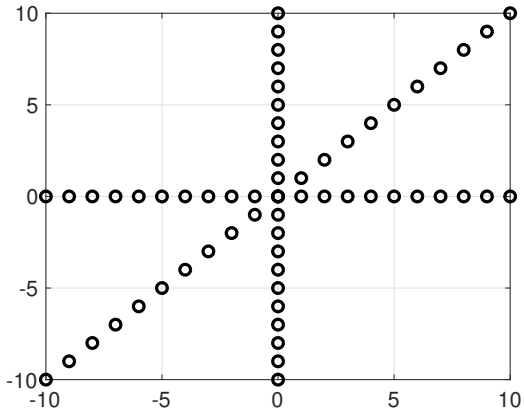
- Open book exam, feel free to use any lecture notes;
- Discussion is not permitted;
- Always give your answer and explain it (guaranteed 5 point for nonempty answer);
- 20 points per problem, totally 110 points (20 * 5 + 10).

0. Write down your name. (10 pts)

1. Consider the following data points (represented by circles) in 2-dimensional space.

- Illustrate the group structure discovered by sparse subspace clustering in the left panel;
- Illustrate the clustering result of single iteration of k-means with $k = 3$ and initial centers $(-10, 0)$, $(0, -10)$, $(10, 10)$ in the right panel.





2. Suppose we have gathered data from n patients as in Table 1, where some entries in the column "Blood Pressure" are missing (represented by the symbol "?"), and other columns are fully observed. Our goal is to estimate these missing values based on the current data matrix. Formulate it as a machine learning problem and state how we can make prediction.

Table 1: Patient data.

	Age	Weight	Height	Gender	Blood Pressure	$\cdots$	Sharp Pain
Patient 1	z_{11}	z_{12}	z_{13}	z_{14}	**?**	$\cdots$	z_{1m}
Patient 2	z_{21}	z_{22}	z_{23}	z_{24}	z_{25}	$\cdots$	z_{2m}
Patient 3	z_{31}	z_{32}	z_{33}	z_{34}	**?**	$\cdots$	z_{3m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Patient n	z_{n1}	z_{n2}	z_{n3}	z_{n4}	z_{n5}	$\cdots$	z_{nm}

3. Now suppose we have another set of data from n patients as in Table 2, where for each column and each row there are some missing entries (represented by the symbol "?"). State when and how we can estimate all the missing values.

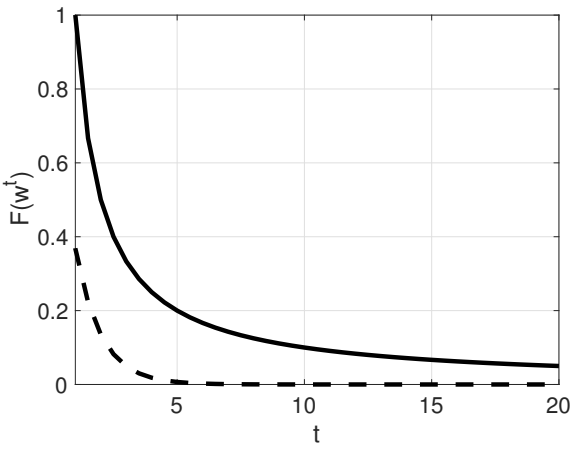
Table 2: Patient data.

	Age	Weight	Height	Gender	Blood Pressure	$\cdots$	Sharp Pain
Patient 1	**?**	z_{12}	z_{13}	z_{14}	**?**	$\cdots$	z_{1m}
Patient 2	**?**	z_{22}	**?**	z_{24}	z_{25}	$\cdots$	**?**
Patient 3	z_{31}	**?**	z_{33}	**?**	**?**	$\cdots$	z_{3m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Patient n	**?**	z_{n2}	z_{n3}	**?**	z_{n5}	$\cdots$	**?**

4. The error type of false positive is defined as follows: an algorithm outputs positive for a sample but in reality its label is negative. In some real-world applications, a learning algorithm should never incur the error of false positive. For example, the face recognition system of a laptop is designed such that it always denies unauthorized access to confidential data. Likewise, a self-driving car shall never recognize a red traffic light as green (but it is fine to classify green as red). Given a set of training samples $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^n \subset \mathbb{R}^d \times \{+1, -1\}$, state a proper formulation which is capable of preventing false positive.

5. Consider two functions $F_1(\boldsymbol{w})$ and $F_2(\boldsymbol{w})$: both of them are strongly convex, but F_1 is smooth and F_2 is non-smooth. Suppose we apply GD to optimize these two functions. The following figure shows two convergence curves: a solid line and a dashed line. One is for $F_1(\boldsymbol{w}^t)$ and another for $F_2(\boldsymbol{w}^t)$.

- Explain which curve may correspond to F_1;
- Plot another possible convergence curve of applying GD to optimize F_2 with a different initial iterate or learning rate.


1
0.8
0.6
0.4
0.2
0
F(w^t)
5
10
15
20
t