EECS 16A Designing Information Devices and Systems I
Summer 2023 Discussion 07D

1. Building a Classifier

We would like to develop a classifier to classify points based on their distance from the origin.

You are presented with the following data. Each data point $\vec{d}_i^{\,T} = [x_i\ y_i]^T$ has the corresponding label $l_i \in \{-1, 1\}$.

x_i	y_i	l_i
-2	1	-1
-1	1	1
1	1	1
2	1	-1

Table 1: *
Labels for data you are classifying

(a) You want to build a model to understand the data. You first consider a linear model, i.e. you want to find $\alpha, \beta, \gamma \in \mathbb{R}$ such that $l_i \approx \alpha x_i + \beta y_i + \gamma$.
Set up a least squares problem to solve for α, β, and γ. If this problem is solvable, solve it, i.e. find the best values for α, β, γ. If it is not solvable, justify why.

UCB EECS 16A, Summer 2023, Discussion 07D, All Rights Reserved. This may not be publicly shared without explicit permission.

(b) You now consider a model with a quadratic term: $l_i \approx \alpha x_i + \beta x_i^2$ with $\alpha, \beta \in \mathbb{R}$. *Read the equation carefully!*

Set up a least squares problem to fit the model to the data. If this problem is solvable, solve it, i.e, find the best values for α, β. If it is not solvable, justify why.

x_i	y_i	l_i
-2	1	-1
-1	1	1
1	1	1
2	1	-1

Table 2: *
Labels for data you are classifying

(c) Finally, you consider the model: $l_i \approx \alpha x_i + \beta x_i^2 + \gamma$, where $\alpha, \beta, \gamma \in \mathbb{R}$. Independent of the work you have done so far, **would you expect this model or the model in part (b) (i.e. $l_i \approx \alpha x_i + \beta x_i^2$) to have a smaller error in fitting the data? Explain why.**

UCB EECS 16A, Summer 2023, Discussion 07D, All Rights Reserved. This may not be publicly shared without explicit permission.

2. Polynomial Fitting

Even though least squares can only be applied to linear systems, it turns out that it can also solve problems with decidedly nonlinear elements. In lecture, you will see an example of fitting to an ellipse. Here, we will fit to a nonlinear polynomial.

Say we know that y is a quartic polynomial in x. This means that we know that y and x are related as follows:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$$

We're also given the following observations, and our goal is to figure out the relationship between x and y:

x	y
0.0	24.0
0.5	6.61
1.0	0.0
1.5	−0.95
2.0	0.07
2.5	0.73
3.0	−0.12
3.5	−0.83
4.0	−0.04
4.5	6.42

(a) What are the unknowns in this question?

(b) Can you write an equation corresponding to the first observation (x_0, y_0), in terms of a_0, a_1, a_2, a_3, and a_4? What does this equation look like? Is it linear in the unknowns?

(c) Now, write a system of equations in terms of a_0, a_1, a_2, a_3, and a_4 using *all of the observations*.

(d) Finally, solve for a_0, a_1, a_2, a_3, and a_4 using IPython or any method you like. You have now found the quartic polynomial that best fits the data!

(e) What if we didn't know the degree of the polynomial? Use the IPython Notebook to explore what happens when we choose a polynomial degree other 4 and explain what you see.

(f) OPTIONAL: Play around with what happens when you add more noise to the data or if you decide to drop data points on the IPython Notebook. Additionally, explore what you see when you change the degree of the polynomial alongside these factors.