Least Squares Method

CS 111: Introduction to Computational Science

Spring 2019 Lecture #9

Ziad Matni, Ph.D.

A I EEE Hacks

WIN THE BATTLE ROYALE TO WIN A SWITCH OR PS4

May 11, 2019 | ESB 1001

ucsbieee.org/ieeehacks

Administrative

- No homework this week!
- Midterm #1 is on Thursday (May 2^{nd})

Midterm Info

- What's on it?
  - Everything (excluding today's least square material)

- What should I bring with me?
  - Pencil, eraser
  - 1 sheet of notes (optional) – double-sided 8.5"x11"
  - Your UCSB ID
  - ***THAT'S IT!!!!***

- What should I NOT bring with me?
  - Any computational device: computer, phone, watch, etc…
  - Any book or other notes

Midterm Info

- We start at 2:00 PM SHARP!
- We have 75 minutes
- You should arrive 10 minutes EARLY
- I definitely WILL be re-seating a lot of you

Lecture Outline

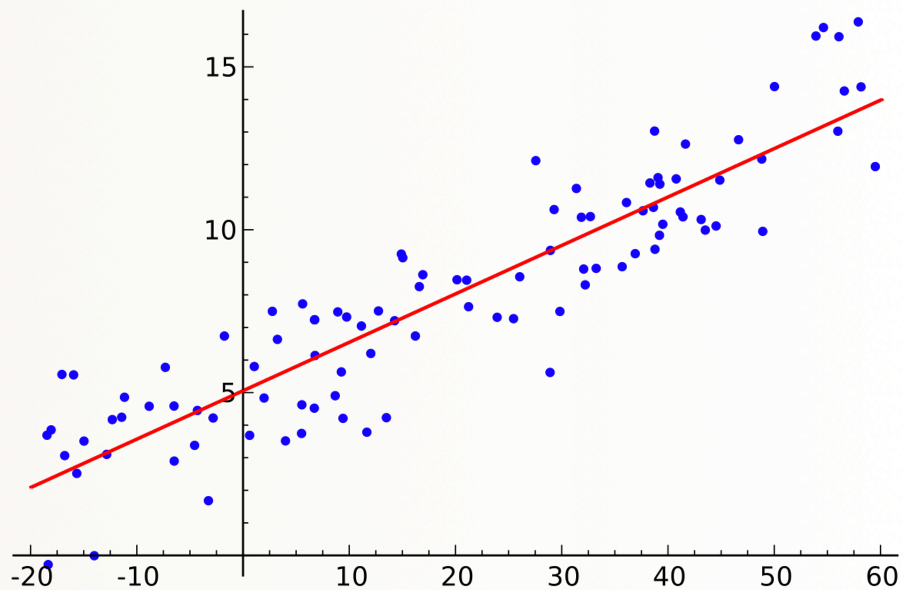
- Least Squares Method

Least Squares Method

A form of mathematical ***regression analysis*** that finds the ***line of best fit*** for a ***set of data***

- Provides a visual demonstration of the relationship between the data points

- Often we refer to the x-axis as the Independent Variables & y-axis as the Dependent Variables



Overdetermined Systems

- We "smooth out" variations in the data by using more data points than needed
  - That is, producing more equations than unknowns
  - This is called an *overdetermined system* (there is no unique solution)

- In a system described with $\mathbf{Ax} = \mathbf{b}$
  - We minimize the distance between the right and left sides
    - i.e. make $||\mathbf{r}|| = ||\mathbf{b} - \mathbf{Ax}||$ as close to zero as we can

- To reflect this lack of exact equality, we often write the system as: $\mathbf{Ax} \cong \mathbf{b}$

Example

- Surveyor wants to determine the heights of 3 hills above some ref. point
- His survey reveals:
  - $x_1 = 1237$ ft.
  - $x_2 = 1941$ ft.
  - $x_3 = 2417$ ft.
- To confirm these, he climbs each one and measures the differences:
  - $x_2 - x_1 = 711$ ft.
  - $x_3 - x_1 = 1177$ ft.
  - $x_3 - x_2 = 475$ ft.
- See blackboard for continuation

Data Fitting Example

- Given data points (t_i, y_i) (i= 1 to m)
- We want to find a vector **x** of parameters that give us the "best fit" to the data by the model function: ***f*(t, x)** as

$$\min \sum_{i=1}^{m} (y_i - f\,(t_i - x))^2$$

Data Fitting Example

- We will assume the model is **linear**, i.e:

$$f(t, x) = x_1 + x_2 t + x_3 t^2 + x_4 t^3 + \ldots + x_n t^{n-1}$$

- So we can express this as a linear system of equations using matrices (see blackboard)

Example Continued using Python Implementation

- Using `npla.lstsq()`

Your To-Dos

- **Study for your midterm! ☺**

</LECTURE>